

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

3M nucleic - nucleic search, using sw model  
Run on: March 10, 2004, 08:38:17 ; Search time 3862.62 Seconds  
(without alignments)  
17011.249 Million cell updates/sec

Title: US-10-084-817-8  
Perfect score: 1516  
Sequence: 1 cttgttttggacatagct.....gctgaataaaatgncgtc 1516

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_on.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rtd.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_higo\_hum.\*  
40: em\_higo\_mus.\*  
41: em\_higo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1389.4	91.6	1399	6	AX440468 Sequence
2	1387.8	91.5	1399	6	AX380797 Sequence
3	1387.8	91.5	1399	6	AX004164 Sequence
4	1387.8	91.5	1399	6	AX281593 Sequence
5	1387.8	91.5	1399	6	AX334637 Sequence
6	1387.8	91.5	1399	6	AX779855 Sequence
7	1387.8	91.5	1399	9	HSLLPCR
8	1375.8	90.8	1398	6	AX004167 Sequence
9	1367.4	90.2	1408	9	BC001275 Homo sapi
10	1364.8	90.0	1376	6	I02238 Sequence 1
11	1364.8	90.0	1376	6	I02432 Sequence 3
12	1364.8	90.0	1376	6	I08414 Sequence 20
13	1357.8	89.6	1385	9	BC035993 Homo sapi
14	1329.8	87.7	2239	9	AK074480 Homo sapi
15	976	64.4	1252	4	AF544227 Equus cab
16	947.6	62.5	1342	4	SSANEXNI
17	938.4	61.9	1335	4	BIANNI
18	917.8	60.5	1327	4	OCU24656
19	913.8	60.3	1379	10	CCLIPP33
20	886.8	58.5	1402	6	AX401909
21	886.8	58.5	1402	10	NMLCI
22	885.2	58.4	1355	10	MMLCIR
23	884.8	58.4	1409	10	BC061710 Rattus no
24	882.2	58.2	1383	10	BC002289 Mus muscu
25	880	58.0	1353	10	RATCALPII
26	880	58.0	1355	6	E01751
27	879.4	58.0	1933	6	E01979
28	877.2	57.9	1358	10	BC004594
29	852	56.2	1330	10	MUSLCI
30	599.4	38.5	1369	5	PGNAXI
31	589.4	38.9	614	6	AX440388
32	566	37.3	567	6	AX885063
33	566	37.3	567	6	BD024673
34	555.4	36.6	1224	5	PGNCAL
35	540.2	35.6	1237	5	BC053786
36	520.4	34.3	550	6	AX197684
37	519.4	34.3	537	6	AX208294
38	457.6	30.2	1409	5	BC064261
39	453.2	29.9	1235	5	AY178793
40	451.6	29.8	1206	5	BC053190
41	446.2	29.4	1337	5	BC054187
42	424.4	28.0	438	6	AX321079
43	422.8	27.9	1544	5	OLANNEXM3
44	362.8	23.9	542	6	BD216773
45	349.6	23.1	1280	5	BC041306

ALIGNMENTS

RESULT 1  
AX440468  
LOCUS AX440468 1399 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 321 from Patent WO0190154.  
ACCESSION AX440468  
VERSION AX440468.1 GI:21665278  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Xu, J., Mitcham, J.L., Harlocker, S.L., Dillon, D.C., Secrist, H.,  
Lodes, M.J., Algate, P.A., Fling, S.P., Mannion, J., Benson, D.R. and  
Carter, D.





121 ATGAGAGCAGGAATATGTTCAAACTGTGAAGTATCCAAAGGTGGTCCGGATCAGCGG 180  
 305 TGAGCCCTATCTCTACCTTCAATCCATCTCGGATGTCGCTGCTTCATCAAGGCCATAA 364  
 181 TGAGCCCTATCTCTACCTTCAATCCATCTCGGATGTCGCTGCTTCATCAAGGCCATAA 240  
 365 TGGTAAAGGTGTGGATGAAGCAACCATCATGTGATCATCTTAACCTAAAGCAACATGCAC 424  
 241 TGGTAAAGGTGTGGATGAAGCAACCATCATGTGATCATCTTAACCTAAAGCAACATGCAC 300  
 425 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGAGAAAGCCCTGGATGAACAC 484  
 301 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGAGAAAGCCCTGGATGAACAC 360  
 485 TGAAGAAAGCCCTTACAGGTCACTTGAAGAGGTGTTTATGCTCTGCTTAAAGCTCCAG 544  
 361 TTAAGAAAGCCCTTACAGGTCACTTGAAGAGGTGTTTATGCTCTGCTTAAAGCTCCAG 420  
 545 CGCAATTTGATGCTGATGAACCTTCTGCTGCTGATGAAGGCTTTGGAACTGATGAAGATA 604  
 421 CGCAATTTGATGCTGATGAACCTTCTGCTGCTGATGAAGGCTTTGGAACTGATGAAGATA 480  
 605 CTCTAATTTGATGCTGATGAACCTTCAAGAACTTAAAGAAATCAGAGACATTAACAGGTCT 664  
 481 CTCTAATTTGATGCTGATGAACCTTCAAGAACTTAAAGAAATCAGAGACATTAACAGGTCT 540  
 665 ACAGAGAGAACTGAAGAGATCTGGCAAGACATTAACCTCAGACACATCTCGAGATT 724  
 541 ACAGAGAGAACTGAAGAGATCTGGCAAGACATTAACCTCAGACACATCTCGAGATT 600  
 725 TTCGAAACGCTTTGCTTCTCTGCTGCTGATGAAGGCTTGAAGGCTTGGTGTGAATG 784  
 601 TTCGAAACGCTTTGCTTCTCTGCTGCTGATGAAGGCTTGAAGGCTTGGTGTGAATG 660  
 785 TGAATCTGCTGATTCAGATGCTGCTGCTGCTGATGAAGGCTTGAAGGCTTGGTGTGAATG 844  
 661 TGAATCTGCTGATTCAGATGCTGCTGCTGCTGATGAAGGCTTGAAGGCTTGGTGTGAATG 720  
 845 CAGAGCTAAACGCTTGAATACCACTTACCAAGCAAGCTATCCCAACTTCGCGAG 904  
 721 CAGAGCTAAACGCTTGAATACCACTTACCAAGCAAGCTATCCCAACTTCGCGAG 780  
 905 TGTTCAGAAATACCAAGTACAGTACGATGATGATGATGATGATGATGATGATGATGATGAT 964  
 781 TGTTCAGAAATACCAAGTACAGTACGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 965 TGAAGAGTGACATTCAGAAATGCTTCAAGCTATCGTGAAGTGGCCCAAGCAACCCAG 1024  
 841 TGAAGAGTGACATTCAGAAATGCTTCAAGCTATCGTGAAGTGGCCCAAGCAACCCAG 900  
 1025 CTTTCTTTGAGAGAGCTTCAATCAAGCAATGAAAGGTGTTGGAACTGCCATAGGCAT 1084  
 901 CTTTCTTTGAGAGAGCTTCAATCAAGCAATGAAAGGTGTTGGAACTGCCATAGGCAT 960  
 1085 TGATCAGGATTTATGTTTCCCGTCTTGAATTTGATGATGATGATGATGATGATGATGATGAT 1144  
 961 TGATCAGGATTTATGTTTCCCGTCTTGAATTTGATGATGATGATGATGATGATGATGATGAT 1020  
 1145 AGAAGAT 1204  
 1021 AGAAGAT 1080  
 1205 AGAAGAT 1264  
 1081 AGAAGAT 1140  
 1265 TGATCAGAGACTTTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1324  
 1141 TGATCAGAGACTTTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1200  
 1325 AACAGGATACAGTGTAGCTACCTACCTGAGAAATATAGCTTTAATCATTTTATTTTATTTTATTTT 1384

1201 AACGAGTTACAGTGTAGCTACCTACCTGCTGAAATATATAGCTTTAAATCATTTTAT 1260  
 1385 ATTATACTCTGTATTAATAGAGATAGTCCATTTTAAATGTTTCCCAACCCATA 1444  
 1261 ATTATACTCTGTATTAATAGAGATAGTCCATTTTAAATGTTTCCCAACCCATA 1320  
 1445 AACCCCTATACAGTGTGTCTAGTAAACATATACATGAGAAAGATGCTATGTAGCTGAAA 1504  
 1321 AACCCCTATACAGTGTGTCTAGTAAACATATACATGAGAAAGATGCTATGTAGCTGAAA 1380  
 1505 TAAATGCGTC 1516  
 1381 TAAATGCGTC 1392

RESULT 3  
 AX004164  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1  
 AUTHORS  
 TITLE  
 JOURNAL  
 GEORGES E. and Wang, Y.  
 P-40/annexin i and related proteins and their role in multidrug  
 resistance  
 Patent: WO 9921980-A 1 06-MAY-1999;  
 GEORGES ELIAS (CA); UNIV MCGILL (CA)  
 FEATURES  
 Location/Qualifiers  
 1..1399  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 75..1115  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC04890.1"  
 /db\_xref="GI:9927715"  
 /db\_xref="REMBL:CAC04890"  
 /translation="MAVSEFLKQAFIENEEQEVVQTVKSKGPGSVSPYPTNP  
 SDVAALHKAIMVGVDEATIIDLTKRNAQRQIKAAAYLOETGKPLDITKKALTG  
 HLEVVALLKTPAOPDADBLAAKMGICGTDBDTLEILASRTNKEIRDIRNRYREL  
 KRLADITSDTSGDFRNALLSLAKGDSRDEKVEDLADSDARALYBAGERRKGTDV  
 NVNTLLTTRSPQLRRVFTQKITKYKSKHDMNKVLDLELKGDIKCLTAIVKCATSKPA  
 FPAEKILHQAMKGVGTHKALIRIMVSRSEIDMNDIKAFYQKMVIGSLCQAILDETQGD  
 YEKILVALCGGN"

Query Match 91.5%; Score 1387.8; DB 6; Length 1399;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

125 AGTGTGAATCTTTCAGAGAGAAATTTCTCTTTAGTTCTTTTGAAGAGGTAGAGATAAG 184  
 1 AGTGTGAATCTTTCAGAGAGAAATTTCTCTTTAGTTCTTTTGAAGAGGTAGAGATAAG 60  
 185 ACATTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAGCAGGCTGTTTATTGAAA 244  
 61 ACATTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAGCAGGCTGTTTATTGAAA 120  
 245 ATGAAGAGCAGGAATATGTTTCAAACTGTGAAGTCAATCCAAAGGTGTCGCGATCAGCGG 304  
 121 ATGAAGAGCAGGAATATGTTTCAAACTGTGAAGTCAATCCAAAGGTGTCGCGATCAGCGG 180  
 305 TGAGCCCTATCTCTACCTTCAATCCATCTCCTCGGATGTCGCTTGCATAAGGCCATAA 364  
 181 TGAGCCCTATCTCTACCTTCAATCCATCTCCTCGGATGTCGCTTGCATAAGGCCATAA 240  
 365 TGGTTAAAGGTGTGATGAAGCAACCATCATTTGACATTTCTTAACCTTAAGCAACATGCAC 424

ORIGIN

Db 241 TGGTTAAAGGTGGTGAAGCAACCATCATTTGACATTTCTAAAGCGAAACAATGCAC 300  
 QY 425 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 484  
 Db 301 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 360  
 QY 485 TGAAGAAAGCCCTTACAGGTCACTGGAGAGGTTGTTTAGCTGTCTGCTAAACCTCCAG 544  
 Db 361 TTAAGAAAGCCCTTACAGGTCACTGGAGAGGTTGTTTAGCTGTCTGCTAAACCTCCAG 420  
 QY 545 CGCAATTTGATGCTGATGAATCTGCTGCCATGAAGGCCCTTGGAACTGATGAAGATA 604  
 Db 421 CGCAATTTGATGCTGATGAATCTGCTGCCATGAAGGCCCTTGGAACTGATGAAGATA 480  
 QY 605 CTCTAATTTGATGCTGATGAATCTGCTGCCATGAAGGCCCTTGGAACTGATGAAGATA 564  
 Db 481 CTCTAATTTGATGCTGATGAATCTGCTGCCATGAAGGCCCTTGGAACTGATGAAGATA 440  
 QY 661 AAGACTTGGCTGATTCAGATGCCAGGCCCTTGTATGAAGCAGGAGAAAGGAAAGGGGA 720  
 QY 845 CAGACGTAACGCTGTTCAATACCACTCTTACCACAGAGCTATCCACAACTTCGCAGAG 904  
 Db 721 CAGACGTAACGCTGTTCAATACCACTCTTACCACAGAGCTATCCACAACTTCGCAGAG 780  
 QY 905 TGTTTCAGAAATACCAAGTACAGTAAGCATGATGAAGTCTGACCAAGTCTGACCTGGAGT 964  
 Db 781 TGTTTCAGAAATACCAAGTACAGTAAGCATGATGAAGTCTGACCAAGTCTGACCTGGAGT 840  
 QY 965 TGAAGGTGACATTCAGAAATGCTCTCAGCTACAGCTATCGTGAAGTGCACCAAGCAACCCAG 1024  
 Db 841 TGAAGGTGACATTCAGAAATGCTCTCAGCTACAGCTATCGTGAAGTGCACCAAGCAACCCAG 900  
 QY 1025 CTTCCTTTCAGAGAGGCTTCATCAGCCATGAAGGTTTGGAACTCGCCATGAAGCAT 1084  
 Db 901 CTTCCTTTCAGAGAGGCTTCATCAGCCATGAAGGTTTGGAACTCGCCATGAAGCAT 960  
 QY 1085 TGATCAGGATTTAGTTTCCCGTTCTGAAATTTGACATGAATGATATCAAAAGCATTTCTATC 1144  
 Db 961 TGATCAGGATTTAGTTTCCCGTTCTGAAATTTGACATGAATGATATCAAAAGCATTTCTATC 1020  
 QY 1145 AGAAGATGATGTTATCTCCCTTTGCCAAGCCATCTCGGATGAACCAAGGAGATGATG 1204  
 Db 1021 AGAAGATGATGTTATCTCCCTTTGCCAAGCCATCTCGGATGAACCAAGGAGATTTATG 1080  
 QY 1205 AGAAATCTCTGCTGCTCTTTGTGAGAGAACTAAACATTCCTTGTGTTCTCAAGCTA 1264  
 Db 1081 AGAAATCTCTGCTGCTCTTTGTGAGAGAACTAAACATTCCTTGTGTTCTCAAGCTA 1140  
 QY 1265 TGATCAGAGACTTTAATATATATTTTCACTCTAATAGCTTAATAGGAAGTTTCTTC 1324  
 Db 1141 TGATCAGAGACTTTAATATATATTTTCACTCTAATAGCTTAATAGGAAGTTTCTTC 1200  
 QY 1325 AACAGGATTCAGTCTGCTACCTGCTGAAATATAGCTTTTAAATCATTTTAT 1384  
 Db 1201 AACAGGATTCAGTCTGCTACCTGCTGAAATATAGCTTTTAAATCATTTTAT 1260  
 QY 1385 ATTAACTCTCTATAATAGAGATGAAGTCCATTTTAAAAATGTTTCCCAACCATTA 1444  
 Db 1261 ATTAACTCTCTATAATAGAGATGAAGTCCATTTTAAAAATGTTTCCCAACCATTA 1320  
 QY 1445 AAACCCCTATACAGTTGTTCTAGTAAACATATACATGAGAAAGTCTATGCTGAGAA 1504

Db 1321 AAACCCCTATACAGTTGTTCTAGTAAACAATACATGAGAAAGATGTTCTATGCTAGTGAAAA 1380  
 QY 1505 TAAATGNCGTG 1516  
 Db 1381 TAAATGACGTG 1392  
 RESULT 4  
 AX281593 1399 bp DNA linear PAT 02-NOV-2001  
 LOCUS Sequence 2 from Patent WO0177389.  
 DEFINITION  
 AX281593  
 ACCESSION  
 AX281593.1 GI:16608844  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 AUTHORS  
 1. Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,  
 Mikita,T. and Tai,J.  
 Genes expressed in foam cell differentiation  
 Patent: WO 0177389-A 2 18-OCT-2001;  
 JOURNAL Incyte Genomics, Inc. (US)  
 FEATURES  
 LOCATION/Qualifiers  
 source  
 1. .1399  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: g34387"  
 ORIGIN  
 Query Match 91.5%; Score 1387.8; DB 6; Length 1399;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1389; Conservative 3; Indels 0; Gaps 0;  
 QY 125 AGTGTGAATCTTCAGAGAAGATTTCTCTTAGTCTTTTCAGAGAAGTAGAGATAAG 184  
 Db 1 AGTGTGAATCTTCAGAGAAGATTTCTCTTAGTCTTTTCAGAGAAGTAGAGATAAG 60  
 QY 185 ACACTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAGCAGGCCCTGGTTATTGAA 244  
 Db 61 ACACTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAGCAGGCCCTGGTTATTGAA 120  
 QY 245 ATGAGAGCAGAAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGATCAGCGG 304  
 Db 121 ATGAGAGCAGAAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGATCAGCGG 180  
 QY 305 TGAGCCCTTATCTACCTTCAATCCATCCTCGATGTGCTGCTTGCATAGGCCATAA 364  
 Db 181 TGAGCCCTTATCTACCTTCAATCCATCCTCGATGTGCTGCTTGCATAGGCCATAA 240  
 QY 365 TGGTTAAAGGTGGTGAAGCAACCATTCATTAAGGAAAGCCCTGGATGAACAC 424  
 Db 241 TGGTTAAAGGTGGTGAAGCAACCATTCATTAAGGAAAGCCCTGGATGAACAC 300  
 QY 425 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 484  
 Db 301 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 360  
 QY 485 TGAAGAAAGCCCTTACAGGTCACTGGAGAGGTTGTTTAGCTGTCTGCTAAACCTCCAG 544  
 Db 361 TTAAGAAAGCCCTTACAGGTCACTGGAGAGGTTGTTTAGCTGTCTGCTAAACCTCCAG 420  
 QY 545 CGCAATTTGATGCTGATGAATCTGCTGCCATGAAGGCCCTTGGAACTGATGAAGATA 604  
 Db 421 CGCAATTTGATGCTGATGAATCTGCTGCCATGAAGGCCCTTGGAACTGATGAAGATA 480  
 QY 605 CTCTAATTTGATGCTGATGAATCTGCTGCCATGAAGGCCCTTGGAACTGATGAAGATA 564  
 Db 481 CTCTAATTTGATGCTGATGAATCTGCTGCCATGAAGGCCCTTGGAACTGATGAAGATA 540  
 QY 665 ACAGAGAGGAACTGAAGAGAGATCTGGCCAAAGACATTAACAGAGGCTCT 724

[illegible]

RESULT	5
LOCUS	AX334637
SEQUENCE	Sequence 5146 from Patent WO0194629.
DEFINITION	AX334637
ACCESSION	AX334637.1 GI:18125356
VERSION	.
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1
REFERENCE	Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., AUTHORS
PATENT	linear PAT 09-JAN-2002

**TITLE** Horrigan, S., Soppet, D.R. and Weaver, Z.  
 Cancer gene determination and therapeutic screening using signature  
 gene sets  
**JOURNAL** Patent: WO 0194629-A 5146 13-DEC-2001;  
 Avalon Pharmaceuticals (US)  
**FEATURES** Location/Qualifiers  
 source 1..1399  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

Query Match	91.5%;	Score 1387.8;	DB 6;	Length 1399;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1389;	Conservative	0;	Mismatches	3;
			Indels	0;
			Gaps	0;
QY	125	AGTGTCAAAATCTTTACAGAGAAGAAATTTCTCTTTTAGTTCTTTTGCAAGAGGCTAGAGATAAAG	184	
DB	1	AGTGTGAAATCTTCAGAGAAGAAATTTCTCTTTTAGTTCTTTTGCAAGAGGCTAGAGATAAAG	60	
QY	185	ACACTTTTTCAAAAATGGCAATGGTATCAGAATTCCTCAAGCAGGCTCGTTTATTGAAA	244	
DB	61	ACACTTTTTCAAAAATGGCAATGGTATCAGAATTCCTCAAGCAGGCTCGTTTATTGAAA	120	
QY	245	ATCAAGCAGCAGGAATATGTTTCAAACCTGTGAAGTCATCCAAAGGTGGTCCCGGATCAGCGG	304	
DB	121	ATGAAGAGCAGGAATATGTTTCAAACCTGTGAAGTCATCCAAAGGTGGTCCCGGATCAGCGG	180	
QY	305	TGAGCCCTATCTCACTCACTCAATCCATCTCTCGGATGTCGTGCTTGTCATAGGCCATATA	364	
DB	181	TGAGCCCTATCTCACTCACTCAATCCATCTCTCGGATGTCGTGCTTGTCATAGGCCATATA	240	
QY	365	TGCTTAAAGGTGGGATGAAGCAACCATCATTCGACATTTCTAACTTAAGCGGAAACAATGCAC	424	
DB	241	TGCTTAAAGGTGGGATGAAGCAACCATCATTCGACATTTCTAACTTAAGCGGAAACAATGCAC	300	
QY	425	AGCGTCAACAGATCAAAAGCAGCATATCTTCAGGAAACAGGAAAGCCCTCGATGAAACAC	484	
DB	301	AGCGTCAACAGATCAAAAGCAGCATATCTTCAGGAAACAGGAAAGCCCTCGATGAAACAC	360	
QY	485	TGAAGAAGCCCTTACAGGTCACCTTGAGCAGGTTGTTTTAGTCTCTCTTAAAACTCTCAG	544	
DB	361	TTAAGAAGCCCTTACAGGTCACCTTGAGCAGGTTGTTTTAGTCTCTCTTAAAACTCTCAG	420	
QY	545	CGCAATTTGATGCTGATGAACCTTCTGTCGCCATGAAGGSCCTTGGAACTGATGAAGATA	604	
DB	421	CGCAATTTGATGCTGATGAACCTTCTGTCGCCATGAAGGSCCTTGGAACTGATGAAGATA	480	
QY	605	CTCTAATTGAGATTTTGGCATCAAGACTTAACAAAGAAATCAGAGACATTAACAGGGTCT	664	
DB	481	CTCTAATTGAGATTTTGGCATCAAGACTTAACAAAGAAATCAGAGACATTAACAGGGTCT	540	
QY	665	ACAGAGAGGAACTGAAGAGAGATCTGCGCCAAAGACATAACCTCAGACACATCTCGAGATT	724	
DB	541	ACAGAGAGGAACTGAAGAGAGATCTGCGCCAAAGACATAACCTCAGACACATCTCGAGATT	600	
QY	725	TTCCGAACGCTTTGCTTTCTCTTGGCTTAAGGCTGACCGATCTGAGGACTTGGCTGTGAATG	784	
DB	601	TTCCGAACGCTTTGCTTTCTCTTGGCTTAAGGCTGACCGATCTGAGGACTTGGCTGTGAATG	660	
QY	785	AAGACTTTGGCTGATTACAGATGCCAGGSCCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA	844	
DB	661	AAGACTTTGGCTGATTACAGATGCCAGGSCCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA	720	
QY	845	CAGACGTAAACGGTGTCTCAATACCATCTTACCCAGACAGAGGTATCCCAACTTCGCACAG	904	
DB	721	CAGACGTAAACGGTGTCTCAATACCATCTTACCCAGACAGGTATCCCAACTTCGCACAG	780	
QY	905	TGTTTTCAGAAATAACCAAGTACAGTAAAGATGACATGAACAAAGTCTCTGACCTGGAGT	964	
DB	781	TGTTTTCAGAAATAACCAAGTACAGTAAAGATGACATGAACAAAGTCTCTGACCTGGAGT	840	
QY	965	TGAAGGTGACATTTGAGAAATGTCCTCACAGCTATCTGGTGAAGTGGCCCAACAGCAACACG	1024	

```

841 TGAAGGTGACATTGAGAAATGCTCACAGCTATCGTGAAGTGGCGCAACAAGCAAAACAG 900
1025 CTTTCCTTCAGAGAGAGCTTCATCAAGCCATGAAAGTGTGGAACTCGCCATAGGCAAT 1084
901 CTTTCCTTCAGAGAGAGCTTCATCAAGCCATGAAAGTGTGGAACTCGCCATAGGCAAT 960
1085 TGATCAGGATTATGGTTTCCGTTCTGAAATGACATGAATGATCAAAAGCATTTCTATC 1144
961 TGATCAGGATTATGGTTTCCGTTCTGAAATGACATGAATGATCAAAAGCATTTCTATC 1020
1145 AGAAGATGATGTAATCTCCCTTGGCCAGCCATCTGGATGAACCAAGAGAGATATG 1204
1021 AGAAGATGATGTAATCTCCCTTGGCCAGCCATCTGGATGAACCAAGAGATATG 1080
1205 AGAAATCTCGTGGTCTCTTGTGGAGGAACTAAACATTTCCCTTGATGGTCTCAAGCTA 1264
1081 AGAAATCTCGTGGTCTCTTGTGGAGGAACTAAACATTTCCCTTGATGGTCTCAAGCTA 1140
1265 TGATCAGAGAGCTTAATATATATTTTCATCTATAAGCTTAATAGGAAGTTCTTC 1324
1141 TGATCAGAGAGCTTAATATATATTTTCATCTATAAGCTTAATAGGAAGTTCTTC 1200
1325 AACAGGATTACAGTGTAGCTACCTACATGCTGAAATAATAGCTTTAAATCATTTTAT 1384
1201 AACAGGATTACAGTGTAGCTACCTACATGCTGAAATAATAGCTTTAAATCATTTTAT 1260
1385 ATTATACTCTGTATAATAGAGATAAGTCCATTTTAAATATGTTTCCCAACCATTA 1444
1261 ATTATACTCTGTATAATAGAGATAAGTCCATTTTAAATATGTTTCCCAACCATTA 1320
1445 AAACCCCTATCAAGTGTCTTAGTAAACAATACATGAGAAAGATGTCTATGAGCTGAAA 1504
1321 AAACCCCTATCAAGTGTCTTAGTAAACAATACATGAGAAAGATGTCTATGAGCTGAAA 1380
1505 TAAATGNCGTC 1516
1381 TAAATGACGTC 1392

RESULT 6
AX779855 Homo sapiens 1399 bp DNA linear PAT 14-JUL-2003
LOCUS AX779855
DEFINITION Sequence 2012 from Patent WO03039443.
ACCESSION AX779855
VERSION AX779855.1 GI:32696849
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Haerlach, T.; Schoch, C.; Kern, W.; Kohlmann, A.; Schnittger, S.;
Dugas, M.; Ellis, R.; Brors, B. and Mergenthaler, S.
TITLE Novel genetic markers for leukemias
JOURNAL Patent: WO 03039443-A 2012 15-MAY-2003;
Deutsches Krebsforschungszentrum (DE);
Ludwig-Maximilian-Universitaet Muenchen (DE);
PD Dr. Dr. (DB); Schoch, Claudia (DB); Kern, Wolfgang (DE)
FEATURES
source
1. 1399
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 91.5%; Score 1387.8; DB 6; Length 1399;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

125 AGTGTGAATCTTCAGAGAGAAATTTCTTTAGTTCTTTCCAGAGAGGTAGAGATAAG 184
1 AGTGTGAATCTTCAGAGAGAAATTTCTTTAGTTCTTTCCAGAGAGGTAGAGATAAG 60
```

```

185 ACCTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCTGGTTATTGAAA 244
61 ACCTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCTGGTTATTGAAA 120
245 ATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGGATCAGCGG 304
121 ATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGGATCAGCGG 180
305 TGAGCCCTATCTTACCTTCAATCCATCTCGGATGTCGCTGCTTCATAGGCCATAA 364
181 TGAGCCCTATCTTACCTTCAATCCATCTCGGATGTCGCTGCTTCATAGGCCATAA 240
365 TGGTTAAAGGTGTGGATGAAGCAACCATCATTTGACATTTCTAACTAAAGCAAAACAATCCAC 424
241 TGGTTAAAGGTGTGGATGAAGCAACCATCATTTGACATTTCTAACTAAAGCAAAACAATCCAC 300
425 AGCGTCAACAGATCAAAAGCAGCATATCTCCAGGAAAACAGGAAGCCCTCGATGAACAC 484
301 AGCGTCAACAGATCAAAAGCAGCATATCTCCAGGAAAACAGGAAGCCCTCGATGAACAC 360
485 TGAAGAAGCCCTTACAGGTCACCTTGAGGAGGTTGTTTGTAGCTCTCTAAAACTCCAG 544
361 TTAAGAAGCCCTTACAGGTCACCTTGAGGAGGTTGTTTGTAGCTCTCTAAAACTCCAG 420
545 CGCAATTTGATGCTGATGAACCTTCTGCTGCCATGAAGGGCCCTTGGAACTGATGAAGATA 604
421 CGCAATTTGATGCTGATGAACCTTCTGCTGCCATGAAGGGCCCTTGGAACTGATGAAGATA 480
605 CTCATTTGAGATTTGGCATCAAGAACTCAAGAACTCAAGAACTCAAGAACTCAAGAACTCAAG 664
481 CTCATTTGAGATTTGGCATCAAGAACTCAAGAACTCAAGAACTCAAGAACTCAAGAACTCAAG 540
665 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAAACCCTCAGACACATCTCGAGATT 724
541 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAAACCCTCAGACACATCTCGAGATT 600
725 TTCGGAACGCTTTGCTTCTTCTGCTTAAGGTGACCGATCTGAGGACTTTGGTGTGATG 784
601 TTCGGAACGCTTTGCTTCTTCTGCTTAAGGTGACCGATCTGAGGACTTTGGTGTGATG 660
785 AAGACTTCGGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGAGAAAGAGAGGGA 844
661 AAGACTTCGGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGAGAAAGAGAGGGA 720
845 CAGACGTAAAGCTGTTCAATACCATCTTACCACCAAGGCTATCCACAACTTCGACAG 904
721 CAGACGTAAAGCTGTTCAATACCATCTTACCACCAAGGCTATCCACAACTTCGACAG 780
905 TGTTCAGAAATACACCAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 964
781 TGTTCAGAAATACACCAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 840
965 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCTGTAAGTGGCCACCAAGCAACACAG 1024
841 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCTGTAAGTGGCCACCAAGCAACACAG 900
1025 CTTTCCTTCAGAGAGCTTCATCAAGCCATGAAAGTGTGGAACTCGCCATAGGCAAT 1084
901 CTTTCCTTCAGAGAGCTTCATCAAGCCATGAAAGTGTGGAACTCGCCATAGGCAAT 960
1085 TGATCAGGATTATGGTTTCCGTTCTGAAATGACATGAATGATCAAAAGCATTTCTATC 1144
961 TGATCAGGATTATGGTTTCCGTTCTGAAATGACATGAATGATCAAAAGCATTTCTATC 1020
1145 AGAAGATGATGTAATCTCCCTTGGCCAGCCATCTGGATGAACCAAGAGAGATATG 1204
1021 AGAAGATGATGTAATCTCCCTTGGCCAGCCATCTGGATGAACCAAGAGATATG 1080
1205 AGAAATCTCGTGGTCTCTTGTGGAGGAACTAAACATTTCCCTTGATGGTCTCAAGCTA 1264
1081 AGAAATCTCGTGGTCTCTTGTGGAGGAACTAAACATTTCCCTTGATGGTCTCAAGCTA 1140
```



QY	1265	TGATCAGAGACGTTTAAATTATATATTTTCATCCTATAAGCTTAAATAGGAAAGTTTCTTC	1324					
DB	1141	TGATCAGAGACGTTTAAATTATATATTTTCATCCTATAAGCTTAAATAGGAAAGTTTCTTC	1200					
QY	1325	AACAGGATTACAGTGCTACCTACATGCTGAAATATATAGCTTTAAATCATTTTTAT	1384					
DB	1201	AACAGGATTACAGTGCTACCTACATGCTGAAATATATAGCTTTAAATCATTTTTAT	1260					
QY	1385	ATTATAACTCTGTATAATAGAGATAAGTCATTTTTTAAAAATGTTTTTCCCAAAACCATTA	1444					
DB	1261	ATTATAACTCTGTATAATAGAGATAAGTCATTTTTTAAAAATGTTTTTCCCAAAACCATTA	1320					
QY	1445	AAACCCCTACACAGTTGTTCTACTGTAAACAATACATGAGAAAGATGCTATGTAGCTGAAAA	1504					
DB	1321	AAACCCCTACACAGTTGTTCTACTGTAAACAATACATGAGAAAGATGCTATGTAGCTGAAAA	1380					
QY	1505	TAAAAATGNCCTC	1516					
DB	1381	TAAAAATGACGTC	1392					
RESULT 8								
AX004167/c		1398 bp	DNA linear PAT 24-AUG-2000					
LOCUS		Sequence 4 from Patent WO9921980.						
DEFINITION		AX004167						
ACCESSION		AX004167.1 GI:9927716						
VERSION		Homo sapiens (human)						
KEYWORDS		Homo sapiens						
SOURCE		Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
ORGANISM		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
REFERENCE		Georges, B. and Wang, Y.						
AUTHORS		P-40/annexin i and related proteins and their role in multidrug						
TITLE		resistance						
JOURNAL		Patent: WO 9921980-A 4 06-MAY-1999;						
FEATURES		GEORGES ELIAS (CA); UNIV MCGILL (CA)						
source	Location/Qualifiers							
	1. .1398							
	/organism="Homo sapiens"							
ORIGIN	/mol_type="unassigned DNA"							
	/db_xref="taxon:9606"							
Query Match 90.8%; Score 1375.8; DB 6; Length 1398;								
Best Local Similarity 99.7%; Pred. No. 0;								
Matches 1388; Conservative 0; Mismatches 3; Indels 1; Gaps 1;								
QY	125	AGTCGAAATCTTCAGAGAGAAATTTCTTTAGTCTTTTGCAAGAGAGGTAGAGATAAAG	184					
DB	1398	AGTCGAAATCTTCAGAGAGAAATTTCTTTAGTCTTTTGCAAGAGAGGTAGAGATAAAG	1339					
QY	185	ACACTTTTCAAAAATGGCAATGATCAGATTCCTCAAGCAGGCTGCTTTATTGAA	244					
DB	1338	-CACTTTTCAAAAATGGCAATGATCAGATTCCTCAAGCAGGCTGCTTTATTGAA	1280					
QY	245	ATGAAGACGAGCAATATGTTCAAACTGTGAAGTCATCAAAAGTGTGTCGGGATCAGCGG	304					
DB	1279	ATGAAGACGAGCAATATGTTCAAACTGTGAAGTCATCAAAAGTGTGTCGGGATCAGCGG	1220					
QY	305	TGAGCCCTATCTTACCTTCAATCGATCCTGGATGTCGCTTGCCTTCGATAGGCCATA	364					
DB	1219	TGAGCCCTATCTTACCTTCAATCGATCCTGGATGTCGCTTGCCTTCGATAGGCCATA	1160					
QY	365	TGGTTTAAAGGTGTGGATGAAGCAACCATCAATTGACATTTCTAACTAAGCGAAACAATGCAC	424					
DB	1159	TGGTTTAAAGGTGTGGATGAAGCAACCATCAATTGACATTTCTAACTAAGCGAAACAATGCAC	1100					
QY	425	ACCGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAACAC	484					
DB	1099	AGCGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAACAC	1040					

QY	485	TTGAAGAAAGCCCTTACAGGTCACCTTGGAGGAGTTGTTTTAGCTCTGCTTAAAAATCTCCAG	544
DB	1039	TTTAAAGAAAGCCCTTACAGGTCACCTTGGAGGAGTTGTTTTAGCTCTGCTTAAAAATCTCCAG	980
QY	545	CGCAATTTTGATGCTGATGAACCTTCGTGCTGCATGAAGGCCCTTGGAACTGATGAAGATA	604
DB	979	CGCAATTTTGATGCTGATGAACCTTCGTGCTGCATGAAGGCCCTTGGAACTGATGAAGATA	920
QY	605	CTCTAAATTGAGATTTTGGCATCAAGAACTTAACAAAGAAATCAGAGACATTAACACGGTCT	664
DB	919	CTCTAAATTGAGATTTTGGCATCAAGAACTTAACAAAGAAATCAGAGACATTAACACGGTCT	860
QY	665	ACAGAGAGAACTGAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAGATT	724
DB	859	ACAGAGAGAACTGAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAGATT	800
QY	725	TTTCGGAAAGCCCTTTGCTTTCTCTTTGCTTAAGGGTGACCGATCTGAGGACTTTTGGTGTCAATG	784
DB	799	TTTCGGAAAGCCCTTTGCTTTCTCTTTGCTTAAGGGTGACCGATCTGAGGACTTTTGGTGTCAATG	740
QY	785	AAGACTTGGCTGATTCAGATGCGAGGCCCTTGTTATGAAGCAGAGAAAGAGAAAGGGA	844
DB	739	AAGACTTGGCTGATTCAGATGCGAGGCCCTTGTTATGAAGCAGAGAAAGAGAAAGGGA	680
QY	845	CAGACGTAAACGTTGTTCAATACCATCTTACACACAGAGACTATCCACAACCTTCGCAGAG	904
DB	679	CAGACGTAAACGTTGTTCAATACCATCTTACACACAGAGACTATCCACAACCTTCGCAGAG	620
QY	905	TGTTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAAGTTCTGGACCTGGAGT	964
DB	619	TGTTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAAGTTCTGGACCTGGAGT	560
QY	965	TGAAAGCGTCATTTGAGAAATGCCTCACAAGCTATCGTGAAGTGCGCCACAAAGCAAAACCAG	1024
DB	559	TGAAAGCGTCATTTGAGAAATGCCTCACAAGCTATCGTGAAGTGCGCCACAAAGCAAAACCAG	500
QY	1025	CTTTCTTTGCAGAGAAAGCTTCATCAAGCCATGAAGAGTGTTGGAACTCGGCCATAAGGCAT	1084
DB	499	CTTTCTTTGCAGAGAAAGCTTCATCAAGCCATGAAGAGTGTTGGAACTCGGCCATAAGGCAT	440
QY	1085	TGATCAGGATTATGGTTTTCCCGTTCTCGAAATTGACATGAATGATATCAAAAGCAATTCCTATC	1144
DB	439	TGATCAGGATTATGGTTTTCCCGTTCTCGAAATTGACATGAATGATATCAAAAGCAATTCCTATC	380
QY	1145	AGAAGATGATGATGATCTCCCTTTGGCCAGGCCATCCTGGATGAACCAAGAGAGAGCTATG	1204
DB	379	AGAAGATGATGATGATCTCCCTTTGGCCAGGCCATCCTGGATGAACCAAGAGAGATTTATG	320
QY	1205	AGAAATCCCTGGTGCTCTTTGTGGAGGAAACTAAACATTCCTCGTGTCTCAAGCTA	1264
DB	319	AGAAATCCCTGGTGCTCTTTGTGGAGGAAACTAAACATTCCTCGTGTCTCAAGCTA	260
QY	1265	TGATCAGAAAGCAATTTAATTATATATTTTCACTCATTAAGCTTAATAGGAAAGTTCTTCTC	1324
DB	259	TGATCAGAAAGCAATTTAATTATATATTTTCACTCATTAAGCTTAATAGGAAAGTTCTTCTC	200
QY	1325	AACAGGATTACAGTGTAGCTTACATGCTGAAAAATATAGCGTTTAAATCAATTTTTTAT	1384
DB	199	AACAGGATTACAGTGTAGCTTACATGCTGAAAAATATAGCGTTTAAATCAATTTTTTAT	140
QY	1385	ATTATATACTCTGTATAATAGAGATAAGTCAATTTTTTAAATAAGTTTTTCCCAAAACCATA	1444
DB	139	ATTATATACTCTGTATAATAGAGATAAGTCAATTTTTTAAATAAGTTTTTCCCAAAACCATA	80
QY	1445	AAACCCCTATACAGTTGTTCTTAGTAAACAATACATGAGAAAGATGTCTATGTAGCTGAAAA	1504
DB	79	AAACCCCTATACAGTTGTTCTTAGTAAACAATACATGAGAAAGATGTCTATGTAGCTGAAAA	20
QY	1505	TAAATGNCGTC	1516
DB	19	TAAATGACGTC	8



RESULT 9	BC001275	1408 bp	mrna	linear	PRI 04-OCT-2003
LOCUS	Homo sapiens annexin A1, mRNA (cDNA clone MGC:5095 IMAGE:3459615), complete cds.				
DEFINITION	BC001275				
ACCESSION	BC001275				
VERSION	1				
KEYWORDS	MGC				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1408)				
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klautner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Tishiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullah,J.S., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Wollay,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villoren,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickinson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinska,M.I., Skalska,U., Smalius,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
MEDLINE	22388257				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 1408)				
AUTHORS	Strausberg,R.				
JOURNAL	Direct Submission				
TITLE	Submitted (11-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILML) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: <a href="http://www-shgc.stanford.edu">http://www-shgc.stanford.edu</a> Contact: (Dickson, Mark) <a href="mailto:mcdpaxil@stanford.edu">mcdpaxil@stanford.edu</a> Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.				
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILML at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAK Plate: 4 Row: 9 Column: 18 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4502100. Location/Qualifiers 1..1408 /organism="Homo sapiens" /db_xref="taxon:9606" /mol_type="mRNA" /clone="MGC:5095 IMAGE:3459615" /tissue_type="cervix carcinoma" /clone_lib="NIH_MGC_12" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 1..1408 /gene="ANXA1"				
source	gene				



QY 627 AAGAACTAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGAACTGAAGAGAGA 586  
DB 481 AAGAACTAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGAACTGAAGAGAGA 540  
QY 687 TCTGGCCAAAGACATAACCTCAGACACATCTGGAGATTTTCGGAAGCTTTTCTCTCT 746  
DB 541 TCTGGCCAAAGACATAACCTCAGACACATCTGGAGATTTTCGGAAGCTTTTCTCTCT 600  
QY 747 TGTAAAGGTGACCGATCTCAGAGACTTGTGTGAATGAAGACTTGTGTGATTCAGATGC 806  
DB 601 TGTAAAGGTGACCGATCTCAGAGACTTGTGTGAATGAAGACTTGTGTGATTCAGATGC 660  
QY 807 CAGGCTCTGTATGAACAGAGAGAAAGAGAAAGGGGACAGAGCTTAAAGCTGTTCATATAC 866  
DB 561 CAGGCTCTGTATGAACAGAGAGAAAGAGAAAGGGGACAGAGCTTAAAGCTGTTCATATAC 720  
QY 867 CATCTTACCAACAGAGCTATCCAACTTGGCAGAGTGTTCAGAAATACCAAGTA 926  
DB 721 CATCTTACCAACAGAGCTATCCAACTTGGCAGAGTGTTCAGAAATACCAAGTA 780  
QY 927 CAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTCGAAAGGTGACATTCAGAAATG 986  
DB 781 CAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTCGAAAGGTGACATTCAGAAATG 840  
QY 987 CCTCAGAGCTATCGTGAAGTGCACCAAGAAACCAAGCTTTCTTTCAGAGAGAGCTTCA 1046  
DB 841 CCTCAGAGCTATCGTGAAGTGCACCAAGAAACCAAGCTTTCTTTCAGAGAGAGCTTCA 900  
QY 1047 TCAGGCCATGAAGGTGTGGAACTCCCATTAAGCCATTGATCAGGATTAATGTTTCCCG 1106  
DB 901 TCAGGCCATGAAGGTGTGGAACTCCCATTAAGCCATTGATCAGGATTAATGTTTCCCG 960  
QY 1107 TTCTGAATTCGACATGAATGATATCAAGCAATCTATCAGAGAGATGATGGTATCTCCCT 1166  
DB 961 TTCTGAATTCGACATGAATGATATCAAGCAATCTATCAGAGAGATGATGGTATCTCCCT 1020  
QY 1167 TTGCCAAGCCATCTGGATGAACCAAGGAGAGATGAGAAATCTCTGGTCTCTTTG 1226  
DB 1021 TTGCCAAGCCATCTGGATGAACCAAGGAGAGATGAGAAATCTCTGGTCTCTTTG 1080  
QY 1227 TGGAGAACTAAACATTTCCCTTGTATGCTCTCAAGCTATGATCAGAGAGCTTTAATATA 1286  
DB 1081 TGGAGAACTAAACATTTCCCTTGTATGCTCTCAAGCTATGATCAGAGAGCTTTAATATA 1140  
QY 1287 TATTTTCATCTTAAGCTTAAATAGGAAAGTTTCTTCAAGGATTAACAGTGTAGCTAC 1346  
DB 1141 TATTTTCATCTTAAGCTTAAATAGGAAAGTTTCTTCAAGGATTAACAGTGTAGCTAC 1200  
QY 1347 CTACATGCTGAAATAATAGCTTTTAAATCATTTTATATATTAATCTGTATATAGAG 1406  
DB 1201 CTACATGCTGAAATAATAGCTTTTAAATCATTTTATATATTAATCTGTATATAGAG 1260  
QY 1407 ATAAGTCAATTTTAAATGTTTTCCTCCCAACATTAATAACCTTATACAGTGTGTTCTA 1466  
DB 1261 ATAAGTCAATTTTAAATGTTTTCCTCCCAACATTAATAACCTTATACAGTGTGTTCTA 1320  
QY 1467 GTAAACAATACATGAGAGAGATGCTATGCTGTAATGAAATATAAATGACGTC 1516  
DB 1321 GTAAACAATACATGAGAGAGATGCTATGCTGTAATGAAATATAAATGACGTC 1370

RESULT 10  
102238  
LOCUS 1376 bp ss-DNA linear PAT 21-MAY-1993  
DEFINITION Sequence 1 from Patent US 4879224.  
ACCESSION 102238  
VERSION 102238.1 GI:270576  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1376)

AUTHORS

Wallner, B.P., Pepinsky, R., Blake, and Garwin, J.L.  
DNA sequences, recombinant DNA molecules and processes for  
producing human phospholipase inhibitor polypeptides  
Patent: US 4879224-A 1 07-NOV-1989;  
Biogen, Inc.; Cambridge, MA

JOURNAL

FEATURES

Location/Qualifiers  
1..1376  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 90.0%; Score 1364.8; DB 6; Length 1376;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1366; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 148 TTTCTCTTTAGTCTTTTCGAAAGAGTAGAGATAAAGACACTTTTCAAAAATGCGCAATG 207  
DB 1 TTTCTCTTTAGTCTTTTCGAAAGAGTAGAGATAAAGACACTTTTCAAAAATGCGCAATG 60  
QY 208 GTATCAGAAATTCCTCAGCAGCGCTGGTTTATTGAAATGAGAGCAGGAGATATGTTCAA 267  
DB 61 GTATCAGAAATTCCTCAGCAGCGCTGGTTTATTGAAATGAGAGCAGGAGATATGTTCAA 120  
QY 268 ACTGTGAAGTCAATCCAAAGGTGGTCCCGGATCAGCGGTGAGCCCTATCTTACCTTCAAT 327  
DB 121 ACTGTGAAGTCAATCCAAAGGTGGTCCCGGATCAGCGGTGAGCCCTATCTTACCTTCAAT 180  
QY 328 CCATCCTCGAGTGTGCTTGCATTAAGGCCATAATGTTTAAAGGTGTGGATGAAGCA 387  
DB 181 CCATCCTCGAGTGTGCTTGCATTAAGGCCATAATGTTTAAAGGTGTGGATGAAGCA 240  
QY 388 ACCATCTTGTGACATTTAACTAAGCGAAACAATGACAGCGTCAACAGATCAAAAGCAGCA 447  
DB 241 ACCATCTTGTGACATTTAACTAAGCGAAACAATGACAGCGTCAACAGATCAAAAGCAGCA 300  
QY 448 TATCTCCAGGAAACAGGAAAGCCCTCTGGAAGAAACATCTGAAGAAAGCCCTTACAGGTAC 507  
DB 301 TATCTCCAGGAAACAGGAAAGCCCTCTGGAAGAAACATCTGAAGAAAGCCCTTACAGGTAC 360  
QY 508 CTTGAGGAGTGTGTTTGTAGTCTCTGCTTAAATACTCCAGCGCAATTTGATGCTGATGAACTT 567  
DB 361 CTTGAGGAGTGTGTTTGTAGTCTCTGCTTAAATACTCCAGCGCAATTTGATGCTGATGAACTT 420  
QY 568 CGTGTGCTCAATGAAGGCTTGGAACTGATGAAGATACTCTAATTTGAGATTTTGGCATCA 627  
DB 421 CGTGTGCTCAATGAAGGCTTGGAACTGATGAAGATACTCTAATTTGAGATTTTGGCATCA 480  
QY 628 AGAATCTAAAGAAATCAGACATTAACAGGCTTACAGAGAGAACTGAAGAGAGAT 687  
DB 481 AGAATCTAAAGAAATCAGACATTAACAGGCTTACAGAGAGAACTGAAGAGAGAT 540  
QY 688 CTGGCCAAAGACATAACCTCAGACACATCTGGAGATTTTCGGAACGCTTGTCTCTCT 747  
DB 541 CTGGCCAAAGACATAACCTCAGACACATCTGGAGATTTTCGGAACGCTTGTCTCTCT 600  
QY 748 GCTAAGGTGACCGATCTGAGGACTTGGTGTGAATGAAGCTTGGCTGATTCAGATGCC 807  
DB 601 GCTAAGGTGACCGATCTGAGGACTTGGTGTGAATGAAGCTTGGCTGATTCAGATGCC 660  
QY 808 AGGGCTTGTATGAAGCAGGAGAAAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 867  
DB 661 AGGGCTTGTATGAAGCAGGAGAAAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
QY 868 ATCTTTACCAACGAGAGCTATCCAACTTCCAGAGTGTTCAGAAATACACCAAGTAC 927  
DB 721 ATCTTTACCAACGAGAGCTATCCAACTTCCAGAGTGTTCAGAAATACACCAAGTAC 780  
QY 928 AGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGGTGACATTCAGAAATGC 987  
DB 781 AGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGGTGACATTCAGAAATGC 840  
QY 988 CTCAAGCTATCTGTAAGTGGCCCAACCAAGCTTTCTTTCAGAGAGAGCTTCAT 1047

841	DB	CTCAGAGCTATCGTGAAGTCGCGCACCAAGCAACACCAAGCTTTCTTTCCAGAGAGCTTCAT	900
1048	QY	CAAGCCATCAAGAGTCTGTGGAATCTGCCAATTAAGGCATTTGATCAGGATTAATGGTTTCCCGT	1107
901	DB	CAAGCCATGAAAGGTGTGTGGAATCTGCCAATTAAGGCATTTGATCAGGATTAATGGTTTCCCGT	960
1108	QY	TCTGAAATTCACATGAATGATATCAAAAGCAATCTATCAGAAGATGATGGTATCTCCCTT	1167
961	DB	TCTGAAATTCACATGAATGATATCAAAAGCAATCTATCAGAAGATGATGGTATCTCCCTT	1020
1168	QY	TGCCAGGCCATCTCGATGAAATCAAAAGGAGAGTATGAGAAAAATCTCTGGTGGCTCTTTGT	1227
1021	DB	TGCCAGGCCATCTCGATGAAATCAAAAGGAGATTAAGAAAAATCTCTGGTGGCTCTTTGT	1080
1228	QY	GGAGGAAATCAAAACATTTCCCTTGATGCTCTCAAGCTATGATCAGAAGACTTTAATATAT	1287
1081	DB	GGAGGAAATCAAAACATTTCCCTTGATGCTCTCAAGCTATGATCAGAAGACTTTAATATAT	1140
1288	QY	ATTTTTCATCTTAAGCTTAAATAGGAAAGTCTTCTCAACAGATTTACAGTGTAGCTACC	1347
1141	DB	ATTTTTCATCTTAAGCTTAAATAGGAAAGTCTTCTCAACAGATTTACAGTGTAGCTACC	1200
1348	QY	TACATGCTGAAAAATATAGCCTTTTAAATCAATTTTATATTAATCTGTATTAATAGAGA	1407
1201	DB	TACATGCTGAAAAATATAGCCTTTTAAATCAATTTTATATTAATCTGTATTAATAGAGA	1360
1408	QY	TAAGTCATATTTTAAAAATGTTTTTCCCAAAACCAATAAAACCCCTATCAAGTTGTTCTAG	1467
1261	DB	TAAGTCATATTTTAAAAATGTTTTTCCCAAAACCAATAAAACCCCTATCAAGTTGTTCTAG	1320
1468	QY	TAAATATACATGAGAAAGATGCTATCTAGCTGAAAAATAAAATGNCCTC	1516
1321	DB	TAAATATACATGAGAAAGATGCTATCTAGCTGAAAAATAAAATGNCCTC	1369

RESULT 11

I02432

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

ORIGIN

Query Match

Best Local Similarity

Matches 1366; Conservative

90.0%; Score 1364.8; DB 6; Length 1376;

99.8%; Pred. No. 0;

0; Mismatches 3; Indels 0; Gaps 0;

148

TTCCTCTTTAGTCTTTCGCAAGAGGTGAGATAAAGACACTTTTTCAAAAATGGCAATG

207

1

TTTCTCTTTAGTCTTTCGCAAGAGGTGAGATAAAGACACTTTTTCAAAAATGGCAATG

60

208

GTATCAGAAATTCCTCAACGAGCCCTGGTTTATTGAAAAATGAAGACAGCAATATGTTCAA

267

61

GTATCAGAAATTCCTCAACGAGCCCTGGTTTATTGAAAAATGAAGACAGCAATATGTTCAA

120

268

ACTGTGAAGTCAATCAAAAGGTGGTCCCGGATCAGCGGTGAGCCCTATCTCTACCTTCAAT

327

121

ACTGTGAAGTCAATCAAAAGGTGGTCCCGGATCAGCGGTGAGCCCTATCTCTACCTTCAAT

180

QY	748	GCTAAGGGTGAACCGATCTGAGGACTTTTGGTGTGAATGAAGACTTCGGCTGATTCAGATGCC	807
Db	601	GCTAAGGGTGAACCGATCTGAGGACTTTTGGTGTGAATGAAGACTTCGGCTGATTCAGATGCC	660
QY	808	AGGCGCTTGTATGAGCAGGAGAAAGGAGAAAGGGGACAGAGCTAAACGTTGTTCAATACC	867
Db	661	AGGCGCTTGTATGAGCAGGAGAAAGGAGAAAGGGGACAGAGCTAAACGTTGTTCAATACC	720
QY	868	ATCCTTACCACCAAGAGCTATCCACAACATTCGACAGAGTGTTCAGAGAAATACACCAAGTAC	927
Db	721	ATCCTTACCACCAAGAGCTATCCACAACATTCGACAGAGTGTTCAGAGAAATACACCAAGTAC	780
QY	928	AGTAAGCATGACATGAACAAGATTCGAGCATTCGAGGTTCAAAAGGTGACATTGAGAGAAATGC	987
Db	781	AGTAAGCATGACATGAACAAGATTCGAGCATTCGAGGTTCAAAAGGTGACATTGAGAGAAATGC	840
QY	988	CTCAGAGCTATCGTGAAGTCCGCCACAAGCAAAACGAGTTCCTTTGACAGAGAGCTTCAT	1047
Db	841	CTCAGAGCTATCGTGAAGTCCGCCACAAGCAAAACGAGTTCCTTTGACAGAGAGCTTCAT	900
QY	1048	CAAGCCATGAAGGTGTTGAACTCCCATTAAGGCATTCATCAGGATTAATGTTTCCCGT	1107
Db	901	CAAGCCATGAAGGTGTTGAACTCCCATTAAGGCATTCATCAGGATTAATGTTTCCCGT	960
QY	1108	TCTGAAATTCGACATGAATGATCAAGGCAATTCATCAGAAGATGATGGTATATCCTT	1167
Db	961	TCTGAAATTCGACATGAATGATCAAGGCAATTCATCAGAAGATGATGGTATATCCTT	1020
QY	1168	TGCCAGCCATTCCTGGATGAAACCAAGGAGAGCTATGAGAGAAATCCTGGTGCTCTTTGT	1227
Db	1021	TGCCAGCCATTCCTGGATGAAACCAAGGAGAGCTATGAGAGAAATCCTGGTGCTCTTTGT	1080
QY	1228	GGAGGAAACTAAACATTCCTTGGATGCTCAAGCTATGATCAGAAGAGCTTTAAATATAT	1287
Db	1081	GGAGGAAACTAAACATTCCTTGGATGCTCAAGCTATGATCAGAAGAGCTTTAAATATAT	1140
QY	1288	ATTTCATCCTTATAGCTTAAATAGGAAAGTTTCTTCAACAGGATTCACAGTGTAGCTACC	1347
Db	1141	ATTTCATCCTTATAGCTTAAATAGGAAAGTTTCTTCAACAGGATTCACAGTGTAGCTACC	1200
QY	1348	TACATGCTGAAATAATAGCGCTTTAAATCAATTTTATATATTAATCTGTATATATAGAGA	1407
Db	1201	TACATGCTGAAATAATAGCGCTTTAAATCAATTTTATATATTAATCTGTATATATAGAGA	1260
QY	1408	TAAATGCTCAATTTTAAATAATGTTTCCCAACCCATATAAAACCCCTATACAAGTTGTTCTAG	1467
Db	1261	TAAATGCTCAATTTTAAATAATGTTTCCCAACCCATATAAAACCCCTATACAAGTTGTTCTAG	1320
QY	1468	TAAATATATGAGAGAGATCTATAGCTGAGAAATAAAATGTCGTC	1516
Db	1321	TAAATATATGAGAGAGATCTATAGCTGAGAAATAAAATGTCGTC	1369
RESULT 13			
BC035993			
LOCUS			
DEFINITION			
Homo sapiens annexin A1, mRNA (cDNA clone MGC:32774 IMAGE:4662939)			
complete cds.			
ACCESSION			
BC035993			
VERSION			
BC035993.1 GI:23958903			
KEYWORDS			
MGC.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
1 (bases 1 to 1385)			
AUTHORS			
Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,			
Klausch,I., S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,			
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,			
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,			
Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,			
PRI 07-OCT-2000			
linear			
MGC:32774 IMAGE:46629393			

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Butcherfield, Y.S., Krzyzinski, M.I., Skalska, U., Smalls, D.E., Butterfield, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

1247732

2 (bases 1 to 1385)

Strausberg, R.

Direct Submission

Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: CLONTECH

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) [mdc@paxil.stanford.edu](mailto:mdc@paxil.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 41 Row: h Column: 9.

Location/Qualifiers

1. 1385

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:32774 IMAGE:4662939"

/issue\_type="Lung"

/clone\_lib="NIH\_MGC\_77"

/lab\_host="DH10B"

/notes="vector: pDNR-LIB"

1. 1385

/gene="ANKA1"

/notes="synonyms: ANK1, LPC1"

/db\_xref="LocusID:301"

/db\_xref="MIM:151690"

75\_1115

/codon\_start=1

/product="annexin I"

/protein\_id="AAH35993.1"

/db\_xref="GI:23958904"

/db\_xref="LocusID:301"

/translation="MANVSEFLKQAWFIENRQEVYQVTKSGKGGPSVSPYPTNPSSDVAALHKAIVKGVDEATIIDILTKRNNRQOIKAALQETGKPLDELTKALNGLSEVLLAKLTQPAQFDADLRAMKGLTDEDTLIEILASRTNKEIFDINRYVEELKPLADITSDTSGDFRNALLSLAKGDSDFVNEADLSDAALYEAERRRKGTDFVNTILTRSPQLGRVFKTKYKSHDMNVLDLELKGDBLKLTAIVFACATSKPAFFAEKHQAKVKGVTGRHKALIRIMVSRSEIDNWDIKAFYQKYGISLCQALDSTKGDYEKILVALCGGN"

213\_407

/notes="annexin I; Region: Annexin. This family of annexins also includes giardin that has been shown to function as an annexin"

/db\_xref="CDD:pfam00191"

misc\_feature 423..623

/notes="annexin; Region: Annexin. This family of annexins also includes giardin that has been shown to function as an annexin"

/db\_xref="CDD:pfam00191"

587..575

/notes="annexin; Region: Annexin. This family of annexins also includes giardin that has been shown to function as an annexin"

/db\_xref="CDD:pfam00191"

900..1100

/notes="annexin; Region: Annexin. This family of annexins also includes giardin that has been shown to function as an annexin"

/db\_xref="CDD:pfam00191"

ORIGIN

Query Match 89.6%; Score 1357.8; DB 9; Length 1385;

Best Local Similarity 98.8%; Pred. No. 2.1e-313;

Matches 1368; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 125 AGTGTGAATCTTCAGAGAAGAAATTTCTTTAGTTCTTTGCAAGAAGTAGAGATAAG 184

Db 1 AGTGTGAATCTTCAGAGAAGAAATTTCTTTAGTTCTTTGCAAGAAGTAGAGATAAG 60

Qy 185 ACACTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCCCTGTTATTGAAA 244

Db 61 ACACTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCCCTGTTATTGAAA 120

Qy 245 ATCAAGCAGCAGGAATATGTTCAACTGTGAAGTATCAAAAGGTGTCCTCCGATCAGGG 304

Db 121 ATCAAGCAGCAGGAATATGTTCAACTGTGAAGTATCAAAAGGTGTCCTCCGATCAGGG 180

Qy 305 TGAGCCCTTATCTCACTTCAATCCATCTCTCGATGTCGCTGCTTGCATTAAGGCATAA 364

Db 181 TGAGCCCTTATCTCACTTCAATCCATCTCTCGATGTCGCTGCTTGCATTAAGGCATAA 240

Qy 365 TGCTTAAAGGTGGAATGAGCAACCATCATTTGCAATCTAACTAAAGCAACATGAC 424

Db 241 TGCTTAAAGGTGGAATGAGCAACCATCATTTGCAATCTAACTAAAGCAACATGAC 300

Qy 425 AGCGTCAACAGATCAAAAGCAGCATATCTCCAGAAACAGGAAAGCCCTGGATGAACAC 484

Db 301 AGCGTCAACAGATCAAAAGCAGCATATCTCCAGAAACAGGAAAGCCCTGGATGAACAC 360

Qy 485 TGAAGAAGCCCTTACAGTCACCTTGAGGAGGTGTTTGTAGCTCTGCTAAAACCTCAG 544

Db 361 TGAAGAAGCCCTTACAGTCACCTTGAGGAGGTGTTTGTAGCTCTGCTAAAACCTCAG 420

Qy 545 CGCAATTTGATCTGATGAATCTCTGCTGCTGATGAAGGCTTTGGAACTGATGAAGATA 604

Db 421 CGCAATTTGATCTGATGAATCTCTGCTGCTGATGAAGGCTTTGGAACTGATGAAGATA 480

Qy 605 CTCTAATTCAGATTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGGGTCT 664

Db 481 CTCTAATTCAGATTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGGGTCT 540

Qy 665 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATTAACCTCAGACATCTCGAGATT 724

Db 541 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATTAACCTCAGACATCTCGAGATT 600

Qy 725 TTCGGAACGCTTTGCTTTCTCTTCTTGAAGGCTGACCGATCTGAGGACTTTGCTGTGAATG 784

Db 601 TTCGGAACGCTTTGCTTTCTCTTCTTGAAGGCTGACCGATCTGAGGACTTTGCTGTGAATG 660

Qy 785 AGACTTGTGCTGATTCAGATGCGAGGCTTTGTTATGAAGCAGGAGAAAGGAGAAAGGGA 844

Db 661 AGACTTGTGCTGATTCAGATGCGAGGCTTTGTTATGAAGCAGGAGAAAGGAGAAAGGGA 720

Qy 845 CAGACTTAAACGCTTTCAATCCATCTTACCCAGAGCTATCCCACTTCGACAG 904

Db 721 CAGACTTAAACGCTTTCAATCCATCTTACCCAGAGCTATCCCACTTCGACAG 780

misc\_feature

905	QY	TCGTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACCAAGTTCGGACCTGGAGT	964
781	DB	TCGTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACCAAGTTCGGACCTGGAGT	840
965	QY	TGAAAGGTGACATCTGAGAAATGCCTCACAGCTATCGTGAAGTGCGCACCAAGCAAAACCCAG	1024
841	DB	TGAAAGGTGACATCTGAGAAATGCCTCACAGCTATCGTGAAGTGCGCACCAAGCAAAACCCAG	900
1025	QY	CTTTCTTCAGAGAGCTTCATCAGGCCATGAAGGTCGTTGGAACTCGCCATAAGGCAT	1084
901	DB	CTTTCTTCAGAGAGCTTCATCAGGCCATGAAGGTCGTTGGAACTCGCCATAAGGCAT	960
1085	QY	TGATCAGGATTAATGGTTTCCCGTTCTGAAATTGACATGAATGATATCAAAAGCATTTCTATC	1144
961	DB	TGATCAGGATTAATGGTTTCCCGTTCTGAAATTGACATGAATGATATCAAAAGCATTTCTATC	1020
1145	QY	AGAAAGATGATGATGATCTCCCTTTGCGAAGCCATCCTGGATGAAACCAAAAGGAGAGTATG	1204
1021	DB	AGAAAGATGATGATGATCTCCCTTTGCGAAGCCATCCTGGATGAAACCAAAAGGAGATTAAG	1080
1205	QY	AGAAATTCCTGGTGGCTCTTTGTGGAGGAACTAAACATTCCTCGATGGTCTCAGCTA	1264
1081	DB	AGAAATTCCTGGTGGCTCTTTGTGGAGGAACTAAACATTCCTCGATGGTCTCAGCTA	1140
1265	QY	TGATCAGAGACTTTAAATTATATATATTTTCATCTCCTATAAGCTTAAATAGGAAAGTTTCTTC	1324
1141	DB	TGATCAGAGACTTTAAATTATATATATTTTCATCTCCTATAAGCTTAAATAGGAAAGTTTCTTC	1200
1325	QY	AACAGATTACAGTGTAGCTACCTACATGCTGAAATATATAGCCTTTAAATCATTTTTTAT	1384
1201	DB	AACAGATTACAGTGTAGCTACCTACATGCTGAAATATATAGCCTTTAAATCATTTTTTAT	1260
1385	QY	ATTATAACTCTGTATATATAGAGATAAGTCCATTTTTTAAATAATGTTTTCCCAAAACCAT	1444
1261	DB	ATTATAACTCTGTATATATAGAGATAAGTCCATTTTTTAAATAATGTTTTCCCAAAACCAT	1320
1445	QY	AAACCCCTATACAGTGTCTCTAGCTTAACAATACATAGAGAAAGATGCTCTCTAGCTGAAAA	1504
1321	DB	AAACCCCTATACAGTGTCTCTAGCTTAACAATACATAGAGAAAGATGCTCTCTAGCTGAAAA	1380
1505	QY	TAATAA 1509	
1381	DB	AAAAA 1385	

RESULT	14
AK074480	
LOCUS	AK074480
DEFINITION	Homo sapiens cDNA FLJ23900 fis, clone LNGI5486, highly similar to ANEXIN I.
ACCESSION	AK074480
VERSION	AK074480.1 GI:18677095
KEYWORDS	Oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirose M., Ohmori Y., Oka T., Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T. and Sugano S.
TITLE	NEDO human cdna sequencing project
JOURNAL	Unpublished 2 (bases 1 to 2339) Sugano S., Suzuki Y., Ota T., Obayashi M., Nishi T., Isogai T., Shibahara T., Tanaka T. and Nakamura Y.
AUTHORS	Direct Submission
TITLE	Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shiokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:ficdca@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286)
JOURNAL	



QY 963 GTTGAAGGTGACATTTGAGAAATGCTTCACAGCATCTCGTGAAGTGGCCACAGCAACC 1022  
DB 1647 GTTGAAGGTGACATTTGAGAAATGCTTCACAGCATCTCGTGAAGTGGCCACAGCAACC 1706  
QY 1023 AGCTTTCTTTGCGAGAGAGCTTCATCAAGCATGAAGTGTGGAGTTCGCGCATAGGC 1082  
DB 1707 AGCTTTCTTTGCGAGAGAGCTTCATCAAGCATGAAGTGTGGAGTTCGCGCATAGGC 1766  
QY 1083 ATTGATCAGGATTATGTTTCCCGTTCTGAAATGACATGAATATCAAAAGCATTTCTA 1142  
DB 1767 ATTGATCAGGATTATGTTTCCCGTTCTGAAATGACATGAATATCAAAAGCATTTCTA 1826  
QY 1143 TCAGAGATGTATGTTATCTCCCTTTGCGAGAGAGCTTCGATGAAGCAAAAGGAGATTA 1202  
DB 1827 TCAGAGATGTATGTTATCTCCCTTTGCGAGAGAGCTTCGATGAAGCAAAAGGAGATTA 1886  
QY 1203 TGAGAAATCTCGTGGCTCTTTGTGAGAGAACTAAACATTCCTTTGATGTCTCAAGC 1262  
DB 1887 TGAGAAATCTCGTGGCTCTTTGTGAGAGAACTAAACATTCCTTTGATGTCTCAAGC 1946  
QY 1263 TATGATCAGAGAGCTTTAATATATATTTTTCATCCTTAAAGCTTAAATAGGAAAGTTTCT 1322  
DB 1947 TATGATCAGAGAGCTTTAATATATATTTTTCATCCTTAAAGCTTAAATAGGAAAGTTTCT 2006  
QY 1323 TCACAGGATTACAGTGTAGCTACCTACATGCTGAAATATAGCTTTTAAATCATTTTT 1382  
DB 2007 TCACAGGATTACAGTGTAGCTACCTACATGCTGAAATATAGCTTTTAAATCATTTTT 2066  
QY 1383 ATATTAACTCTGTAATAGAGATAGTCCATTTTAAATAGTTCCTCCCAAGCA 1442  
DB 2067 ATATTAACTCTGTAATAGAGATAGTCCATTTTAAATAGTTCCTCCCAAGCA 2126  
QY 1443 TAAACCTTACAGTGTGTTGTAGTAACATACATGAGAAAGATGTCTATGTAGCTGAA 1502  
DB 2127 TAAACCTTACAGTGTGTTGTAGTAACATACATGAGAAAGATGTCTATGTAGCTGAA 2186  
QY 1503 AATTAATGTCGTC 1516  
DB 2187 AATTAATGTCGTC 2200

RESULT 15  
AF544227  
LOCUS AF544227 1252 bp mRNA linear MAM 03-OCT-2002  
DEFINITION Equus caballus lipocortin-1 mRNA, complete cds.  
ACCESSION AF544227  
VERSION AF544227.1 GI:23477705  
KEYWORDS  
SOURCE Equus caballus (horse)  
ORGANISM Equus caballus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
REFERENCE 1 (bases 1 to 1252)  
Bryant, C.E., Allen, A. and Maskell, D.J.  
Cloning of equine lipocortin-1 and its full cDNA sequence  
JOURNAL Unpublished  
AUTHORS Bryant, C.E., Allen, A. and Maskell, D.J.  
TITLE Direct Submission  
REFERENCE 2 (bases 1 to 1252)  
Bryant, C.E., Allen, A. and Maskell, D.J.  
Submitted (09-SEP-2002) Clinical Veterinary Medicine, University of  
Cambridge, Madingley Road, Cambridge CB10 1B, UK  
FEATURES  
source location/Qualifiers  
1..1252  
/organism="Equus caballus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9796"  
168..1208  
/note="annexin 1"  
/codon\_start=1  
/product="lipocortin-1"  
/protein\_id="AA034819.1"  
/db\_xref="GI:23477706"

CDS

/translation="MSMVSAFLKQAWFIENEBOEYIKAVKSGKGGPSAVSPSPFNP  
SDVDALHKAITVKGVDKATITLITKNNARQOIKAALQEKGPDEALKKALTC  
HLEDVALAKITPAPDFQADELRANKGLGTEDDTLIELTSRTNKEIREINRVVREEL  
KRDIAKDJTSDTSGDFQKALLSLAKGSDSEDFGNDLADADRALVEAGRRKGTGV  
NFTNLTITRSYPHLRRUPQMTYKSKHDMKVLDEMKGVENCFTAIIVKCAKSKPM  
FPAKLNAMKAGTRDKILIRIMVSRSEVDNNDIKACYQKLYGLISLCCQALIDETKGD  
YKELVALCGRD"

Query Match 64.4%; Score 976; DB 4; Length 1252;  
Best Local Similarity 88.3%; Pred. No. 2e-222;  
Matches 1072; Conservative 0; Mismatches 140; Indels 2; Gaps 1;

QY 68 GCTAGGTGGCTTCCTTTAAATCCTATATAATCAGAAGCCCAAGCTTCCTCCAGT 127  
DB 37 GCTGGGTGGCTTCCTTTAAATCCTATATAATCAGAAGCCCAAGCTTCCTCCAGT 96  
QY 128 GTGAATCTTTCAGAGAAGAAATTTCTTTAGTTCCTTTGCAAGAAGGTAGAGATAAGACA 187  
DB 97 GTGAATCTTTCAGAGAAGAAATTTCTTTAGTTCCTTTGCAAGAAGGTAGAGATAAGACA 156  
QY 188 CTTTTCRAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGCCCTGGTTTATTGAAATG 247  
DB 157 CTTTTCRAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGCCCTGGTTTATTGAAATG 216  
QY 248 AAGAGCAGGAATATGTTCAAACTGTGAAGTCTATCAAAAGGTGGTCCCGATCAGCGGTGA 307  
DB 217 AAGAGCAGGAATATGTTCAAACTGTGAAGTCTATCAAAAGGTGGTCCCGATCAGCGGTGA 276  
QY 308 GCCCTATCTTACCTTCATTCATTCCTCGGATGTCCTCGCTTGCATAGGCCATATGG 367  
DB 277 GCCCTATCTTACCTTCATTCATTCCTCGGATGTCCTCGCTTGCATAGGCCATATGG 336  
QY 368 TTAAGGTGTGGATGAAGCAACCATTCATTCATTAAGTGAAGCAACATTCACAGC 427  
DB 337 TGAAGGTGTGGATGAAGCAACCATTCATTCATTAAGTGAAGCAACATTCACAGC 396  
QY 428 GTCAACAGATCAAGCAGCATATCTCCAGAAACAGGAAGCCCTGGATGAACACTGA 487  
DB 397 GTCAACAGATCAAGCAGCATATCTCCAGAAACAGGAAGCCCTGGATGAACACTGA 456  
QY 488 AGAAGCCCTTACAGGTCACTTGAAGAGTGTGTTTGTAGCTCTGCTAAAACTCCAGCGC 547  
DB 457 AGAAGCCCTTACAGGTCACTTGAAGAGTGTGTTTGTAGCTCTGCTAAAACTCCAGCGC 516  
QY 548 AATTGATCTGATGAACTTCGTGCTGCATGAGCGCTTGGAACTGATGAGATATGC 607  
DB 517 GGTTCGATCTGATGAACTTCGTGCTGCATGAGCGCTTGGAACTGATGAGATATGC 576  
QY 608 TAATTCAGATTTTGGCATCAAGAACTAAAGAAATCAGAGACATTAACAGGCTTACA 667  
DB 577 TGATTGAAATTTTGACATCAAGAACTAAAGAAATTAAGGAAATTAACAGAGTCTATA 636  
QY 668 GAGAGAACTGAAGAGAGATCTGGCCCAAGACATTAACCTCAGACACATCTCGAGATTTTC 727  
DB 637 GAGAGAACTGAAGAGAGATCTGGCCCAAGACATTAACCTCAGACACATCTCGAGATTTTC 696  
QY 728 GGAACTGCTTTCTCTTGTAAAGGTGACCGATCTGAGGACTTTTGGTGTGAATGAAG 787  
DB 697 AGAAGCTTTGCTTTCTCTTGTAAAGGTGACCGATCTGAGGACTTTTGGTGTGAATGAAG 756  
QY 788 ACTTGGCTGATTCAGATCCCGGCTTGTATGAGCAGGAGAGAGAGAGAGAGAGAGAG 847  
DB 757 ACTTGGCTGATTCAGATCCCGGCTTGTATGAGCAGGAGAGAGAGAGAGAGAGAGAGAG 816  
QY 848 AGCTAAACGTGTTCAATACCATTCCTTACCACAGAGCTATCCCAACTTCGCGAGAGTGT 907  
DB 817 ATGTGAACGTGTTCAATACCATTCCTTACCACAGAGCTATCCCAACTTCGCGAGAGTGT 876  
QY 908 TTCAGAAATACACCAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATG 967  
DB 877 TTCAGATGTACACCAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATG 936

Qy	968	AAGTGCACATTGAGAAATGCTCACAGCTATCGTGAAGTGCOCCTCAAGCAAAACCAAGCTT	1027
Db	937	AGGGGACGTCGAGAACTGCTTCACAGCTATTCTGAAGTGTGCGACAAGCAAAACCAATGT	996
Qy	1028	TCCTTCAGAGAGCTTCATCAAGCCATGAAAGGTGTTGGAACTGCCATTAAGGCATTGA	1087
Db	997	TCCTTCGAGAGAGCTTCATCAAGCCATGAAAGGTGTTGGAACTGCCATTAAGGCATTGA	1056
Qy	1088	TCAGGATTATGGTTTCCCGTTCTTGAATTCACATGAATGATATCAAGCATTTCTATCAGA	1147
Db	1057	TCAGGATTATGGTTTCCCGTTCTTGAATTCACATGAATGATATCAAGCATTTCTATCAGA	1116
Qy	1148	AGATGATGATATCTCCCTTTGCCAAGCCATCCTGGATGAAACCAAGGAGATGAGA	1207
Db	1117	AGTGTATGATATCTCACTCTGCAAGCCATCCTGGATGAAACCAAGGAGATTTACGAAA	1176
Qy	1208	AAATCCTGGTGGCTCTTTGTGGAGGAACTAAACATTCCCTTTGATGGTCTCAAGC--TAT	1265
Db	1177	AGATCCTGGTGGCTCTCTGTGGAGAGACTAAACATTCCCTTTCAAGCATTTT	1236
Qy	1266	GATCAGAAAGACTTT	1279
Db	1237	CATCAGAAAGACTTT	1250

Search completed: March 11, 2004, 03:35:30  
Job time : 3888.73 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

DM nucleic - nucleic search, using sw model  
Run on: March 10, 2004, 08:01:57 ; Search time 368.965 Seconds  
(without alignments)  
17454.974 Million cell updates/sec

Title: US-10-084-817-8  
Perfect score: 1516  
Sequence: 1 cttgttttggacatagct.....gctgaataataatcgcgc 1516

Scoring table: IDENTITY NUC  
Gap 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_23Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1515	99.9	1516	6	ABs62753 Prostate
2	1515	99.9	1516	7	ABx77579 Different
3	1497.4	98.8	1517	4	ABA09124 Human lip
4	1495.6	98.7	2334	4	Aah72822 Human cer
5	1495.6	98.7	2334	4	Aah72850 Human cer
6	1389.4	91.5	1399	6	ABK09784 Human ova
7	1387.8	91.5	1399	2	AAX57357 Human p-4
8	1387.8	91.5	1399	6	ABl66809 Lung canc
9	1387.8	91.5	1399	6	AAs94747 Human DNA
10	1387.8	91.5	1399	6	ABK84511 Human cDN
11	1387.8	91.5	1399	6	ABv78074 Hypoxia-r
12	1379.4	91.0	1431	3	AAC77652 Human can
13	1367.4	90.2	1377	9	ADD18446 Human pro
14	1364.8	90.0	1376	1	AAN60555 Sequence
15	1364.8	90.0	1376	2	AAQ23216 Lambda LC
16	1358.4	89.6	1376	2	AAQ05809 Human lip
17	1309	86.3	1325	1	AAN60556 Sequence
18	1262.2	83.3	1432	6	ABZ11387 Human pol
19	886.8	58.5	1402	6	ABK63678 Rat seque
20	886.8	58.5	1402	7	ABT41862 Toxicity
21	886.8	58.5	1402	9	ADB56147 Toxicity
22	880	58.0	1355	1	AAN82025 Fragment
23	879.4	58.0	1933	1	AAN90120 Human lip

24	772.2	50.9	1211	6	ABZ11388
25	652.8	43.1	730	6	ABZ11386
26	589.4	38.9	614	6	ABK09704 Human ova
27	566	37.3	567	3	AAC00928 Human sec
28	520.4	34.3	550	5	Aah82515 Human ova
29	519.4	34.3	537	4	AAs23953 Human ova
30	488.6	32.2	596	6	ABG59522 Human col
31	478.4	31.6	498	6	ABG60100 Human col
32	459.4	30.3	461	6	ABv87296 Human col
33	457.8	30.2	461	6	ABv87296 Human col
34	444	29.3	683	6	ABQ54881 Human ova
35	433	28.6	434	8	ACH29835 Human tes
36	426.4	28.1	485	8	ACH20064 Human adu
37	424.4	28.0	438	6	AAs61555 Lung smal
38	387.6	25.6	505	8	ACH37092 Human end
39	362.8	23.9	542	2	AAX17442 Human gen
40	349.4	23.0	432	8	ACH19886 Human adu
41	339.8	22.4	966	6	ABN74742 Bovine em
42	332.8	22.0	1001	3	AAC57452 Arachidon
43	332.8	22.0	1563	4	AAI99227 Human exc
44	332.8	22.0	1563	5	AAI63577 Human kid
45	319.4	21.1	1362	3	AAA64662 cDNA enco

ALIGNMENTS

RESULT 1  
ABS62753  
ID ABS62753 standard; cDNA; 1516 BP.  
XX  
AC ABS62753;  
XX  
DT 05-NOV-2002 (first entry)  
XX  
DE Prostate adenocarcinoma associated cDNA #28.  
XX  
KW Prostate cancer; differential expression; cancer diagnosis;  
KW cancer treatment; cancer monitoring; prostate adenocarcinoma; gene; ss.  
XX Homo sapiens.  
XX  
PN US2002119463-A1.  
XX  
PD 29-AUG-2002.  
XX  
PF 30-JUL-2001; 2001US-00919172.  
XX  
PR 28-JUL-2000; 2000US-0222469P.  
XX  
FA (FARI/) FARIS M.  
XX (TURN/) TURNER C M.  
XX Faris M, Turner CM;  
XX WPI: 2002-608155/65.  
XX P-PSDB; ABG77176.  
XX  
PT New composition, useful for treating and diagnosing prostate cancer,  
PT comprises human cDNAs that are differentially expressed in prostate  
XX cancer.  
XX  
PS Claim 1; Page 66-67; 147pp; English.  
XX  
CC The invention describes a composition of a number of human cDNAs that are  
CC differentially expressed in prostate cancer. The composition of the  
CC invention useful for a high-throughput method for detecting differential  
CC expression of cDNAs in a nucleic acid containing sample comprising  
CC amplifying the nucleic acids of the sample, hybridizing the composition  
CC with nucleic acids of the sample, detecting the hybridisation complexes  
CC and comparing the complexes with those of a standard, where differences  
CC indicates differential expression. The sample is from a subject with  
CC prostate cancer and comparison with a standard defines an early, mid or



05-OCT-2000; 2000US-0238331P.

(CHEN/) CHEN H.

Chen H;

WPI; 2003-182653/18.

New cDNAs, which are differentially expressed in (metastatic) breast cancer useful for diagnosing or staging, breast cancer, or for monitoring the treatment of breast cancer in an individual.

Claim 1; SEQ ID NO 92; 30pp; English.

The invention describes a combination of cDNAs (designated BC-cDNAs), which are differentially expressed in breast cancer. The combination includes 152 cDNA sequences, or their complements. The protein encoded by any of these BC-cDNAs is useful for screening several molecules or compounds to identify at least one ligand that specifically binds the protein, producing or preparing polyclonal or monoclonal antibodies, or purifying antibodies from a sample. The antibodies, which specifically bind the protein differentially expressed in breast cancer is useful for detecting the expression of a protein in a sample. The BC-cDNAs are also useful for diagnosing, monitoring the treatment of, or staging, breast cancer. This sequence represents a differentially expressed breast cancer associated cDNA. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at seqdata.uspto.gov/sequence.html?DocID=20020156263

Sequence 1516 BP; 493 A; 301 C; 322 G; 399 T; 0 U; 1 Other;

Query Match 99.9%; Score 1515; DB 7; Length 1516;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CTTTGTTTGGACATAGCTGAGCCATGACTTCAACAGAGGCGCCATTAACCT	60
DB	1	CTTTGTTTGGACATAGCTGAGCCATGACTTCAACAGAGGCGCCATTAACCT	60
QY	61	TCCTGTTGCTAGTGTGGTCTCTTAAATCCTATAAAATCAGAAGCCCAAGTCTCCAC	120
DB	61	TCCTGTTGCTAGTGTGGTCTCTTAAATCCTATAAAATCAGAAGCCCAAGTCTCCAC	120
QY	121	TGCCAGTGTGAATCTTTCAGAGAAATTTCTTTAGTTCTTTTCAAGAGGTAGAGAT	180
DB	121	TGCCAGTGTGAATCTTTCAGAGAAATTTCTTTAGTTCTTTTCAAGAGGTAGAGAT	180
QY	181	AAAGACATTTTCAAAATGCGAATGGTATCAGAAATCCTCAAGCAGCCCTGGTTATT	240
DB	181	AAAGACATTTTCAAAATGCGAATGGTATCAGAAATCCTCAAGCAGCCCTGGTTATT	240
QY	241	GAAATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCATCAAAAGGTGTCCTGGATCA	300
DB	241	GAAATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCATCAAAAGGTGTCCTGGATCA	300
QY	301	CGCGTGAGCCCTTATCCTACCTTCATCTCGATGTCCTCGATGTCCTCGATGTCCTCGATG	360
DB	301	CGCGTGAGCCCTTATCCTACCTTCATCTCGATGTCCTCGATGTCCTCGATGTCCTCGATG	360
QY	361	ATAATGGTTAAAGGTGTGGATGAAGCAACCATTCATTCATTAAGCGAAACAAAT	420
DB	361	ATAATGGTTAAAGGTGTGGATGAAGCAACCATTCATTCATTAAGCGAAACAAAT	420
QY	421	GCACAGGTCAACAGATCAAGCAGCATATCTCCAGGAACAGGAGCCCTGGATGAA	480
DB	421	GCACAGGTCAACAGATCAAGCAGCATATCTCCAGGAACAGGAGCCCTGGATGAA	480
QY	481	ACACTGAAGAAAGCCCTTACAGGTCACTTGTAGGAGGTTGTTTGTAGTCTGTAAAACT	540
DB	481	ACACTGAAGAAAGCCCTTACAGGTCACTTGTAGGAGGTTGTTTGTAGTCTGTAAAACT	540
QY	541	CCAGCGCAATTTGATGCTGATGAATCTCGTGTGCCATGAGGGCCCTTGGAACTGATGAA	600

RESULT 3

ABA09124/c

ID ABA09124 standard; cDNA; 1517 BP.

XX

AC ABA09124;



```
QY 917 ACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGCTTGAAAGGTGACA 976
DB 617 ACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGCTTGAAAGGTGACA 558
QY 977 TTGAGAAATGCTCAGAGCTATCGTAAGTGGCCGACAAAGCAACAGCTTCTTTGCGAG 1036
DB 557 TTGAGAAATGCTCAGAGCTATCGTAAGTGGCCGACAAAGCAACAGCTTCTTTGCGAG 498
QY 1037 AGAAGCTTCATCAAGCCATGAAGGTGTGGAATCGCCATAAGGCAATGATCAGGATTA 1096
DB 497 AGAAGCTTCATCAAGCCATGAAGGTGTGGAATCGCCATAAGGCAATGATCAGGATTA 438
QY 1097 TGGTTTCCGTTCTGAATTTGACATGAATGATATCAAGCAATTCATCAGAAAGTATG 1156
DB 437 TGGTTTCCGTTCTGAATTTGACATGAATGATATCAAGCAATTCATCAGAAAGTATG 378
QY 1157 GTATCTCCCTTTGCCAAGCCATCTGGATGAACCAAGAGGAGATGATGAGAAATCTCTGG 1216
DB 377 GTATCTCCCTTTGCCAAGCCATCTGGATGAACCAAGAGGAGATGATGAGAAATCTCTGG 318
QY 1217 TGGCTCTTTGTGGAGGAACATAAATCCCTTGATGTCTCAAGCTATGATCAGAAAGAC 1276
DB 317 TGGCTCTTTGTGGAGGAACATAAATCCCTTGATGTCTCAAGCTATGATCAGAAAGAC 258
QY 1277 TTTAATTATATATTTTCATCTTAAGCTTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1336
DB 257 TTTAATTATATATTTTCATCTTAAGCTTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 198
QY 1337 GTGTAGCTACCTACATGCTGAAAATATAGCCTTTAAATCAATTTTATATATATTAATCTG 1396
DB 197 GTGTAGCTACCTACATGCTGAAAATATAGCCTTTAAATCAATTTTATATATATTAATCTG 138
QY 1397 TATAATAGAGTAAGTCAATTTTAAATGTTTTCCTCCAAACCAATAAACCCCTATACA 1456
DB 137 TATAATAGAGTAAGTCAATTTTAAATGTTTTCCTCCAAACCAATAAACCCCTATACA 78
QY 1457 AGTTGCTTCTAGTAACAATACATGAGAAAGATGCTCTATAGCTGAAATATAAATGTCGTC 1516
DB 77 AGTTGCTTCTAGTAACAATACATGAGAAAGATGCTCTATAGCTGAAATATAAATGACGTC 18

RESULT 4
AAH72822
ID AAH72822 standard; cDNA; 2334 BP.
XX
AC AAH72822;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 4096.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
XX
PR 21-DEC-1999; 99US-017350P.
XX
PR 14-MAR-2000; 2000US-0189315P.
XX
PR 12-MAY-2000; 2000US-0203791P.
XX
PR 09-JUN-2000; 2000US-0210600P.
XX
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX
```

```
DR WPI; 2001-375006/39.
XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
XX Claim 1; Page 870-871; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
XX useful for gene therapy
SQ Sequence 2334 BP; 774 A; 570 C; 510 G; 480 T; 0 U; 0 Other;
Query Match 98.7%; Score 1495.6; DB 4; Length 2334;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1498; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 14 CATAGCTGAGCCATGTAATCTTCAACAGAGGAGGAGGCAATTAATACTTCTGCTGCTAGG 73
DB 16 CGTGGGTGAGCCATGTAATCTTCAACAGAGGAGGAGGCAATTAATACTTCTGCTGCTAGG 75
QY 74 TGTGGCTTCTTTAAATCTTATAAAATCAGAAAGCCCAAGTCTCCACTGCCAGTGTGAAA 133
DB 76 TGTGGCTTCTTTAAATCTTATAAAATCAGAAAGCCCAAGTCTCCACTGCCAGTGTGAAA 135
QY 134 TCTTCAGAGAAGAAATTTCTTTAGTTTCTTTGTAAGAGAGGTAGATTAAGACACATTTT 193
DB 136 TCTTCAGAGAAGAAATTTCTTTAGTTTCTTTGTAAGAGAGGTAGATTAAGACACATTTT 195
QY 194 CAAATATGCAATGGTATCAGATTCCTCAACAGAGCCCTGGTTTATTGAAATGAAGAGC 253
DB 196 CAAATATGCAATGGTATCAGATTCCTCAACAGAGCCCTGGTTTATTGAAATGAAGAGC 255
QY 254 AGGAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGTCCTCCGATCAGCGGTAGCCCT 313
DB 256 AGGAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGTCCTCCGATCAGCGGTAGCCCT 315
QY 314 ATCTACCTTCAATCCATCTCGGATGTCGCTTGCATTAAGGCCATAATGTTTAAAG 373
DB 316 ATCTACCTTCAATCCATCTCGGATGTCGCTTGCATTAAGGCCATAATGTTTAAAG 375
QY 374 GTCTGGATGAAGCAACCATCATTTGACATTTAACTAAGCGAAACAATGCACAGCGTCAAC 433
DB 376 GTCTGGATGAAGCAACCATCATTTGACATTTAACTAAGCGAAACAATGCACAGCGTCAAC 435
QY 434 AGATCAAGAGCATATCTCCAGAAACAGGAGCCCTGGATGAACACACTGAAGAAAG 493
DB 436 AGATCAAGAGCATATCTCCAGAAACAGGAGCCCTGGATGAACACACTGAAGAAAG 495
QY 494 CCCTTACAGGTCACTTGGAGAGGTTGTTTATAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 553
DB 496 CCCTTACAGGTCACTTGGAGAGGTTGTTTATAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 555
QY 554 ATGCTGATGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
DB 556 ATGCTGATGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615
QY 614 AGATTTTGGCATCAAGAACTTAACAAAGAAATCAGAGACATTAACAGGGTCTCAGAGAGG 673
DB 616 AGATTTTGGCATCAAGAACTTAACAAAGAAATCAGAGACATTAACAGGGTCTCAGAGAGG 675
QY 674 AACTGAAGAGAGATCTGGGCCAAGACATAACCTCAGACACATCTGGAGATTTTCGGAAG 733
DB 676 AACTGAAGAGAGATCTGGGCCAAGACATAACCTCAGACACATCTGGAGATTTTCGGAAG 735
QY 734 CTTTGTCTTCTCTTGTCTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTGG 793
DB 736 CTTTGTCTTCTCTTGTCTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTGG 795
```

```

QY 794 CTGATTCAGATGCGAGGCTTGTATGAAGCAGAGAGAAAGGAGGACAGACGTAA 853
Db 796 CTGATTCAGATGCGAGGCTTGTATGAAGCAGAGAGAAAGGAGGACAGACGTAA 855
QY 854 ACCTGTTCAATACATCTTACACAGAGAGCTATCCAACTTCGCAGAGTGTTCAGA 913
Db 856 ACCTGTTCAATACATCTTACACAGAGAGCTATCCAACTTCGCAGAGTGTTCAGA 915
QY 914 AATACACCAAGTACAGTAAGCATGATGAAAGTTCCTGGACCTGGAGTGTGAAGGTG 973
Db 916 AATACACCAAGTACAGTAAGCATGATGAAAGTTCCTGGACCTGGAGTGTGAAGGTG 975
QY 974 ACATGTAGAAATGCTTACAGCTATCGTGAAGTGCAGCAAGCAAAACAGCTTTCTTG 1033
Db 976 ACATGTAGAAATGCTTACAGCTATCGTGAAGTGCAGCAAGCAAAACAGCTTTCTTG 1035
QY 1034 CAGAGAGCTTCATCAAGCCATGAAGGCTGTGGAAGTGCCTGCAATAGGCAATGATCAGA 1093
Db 1036 CAGAGAGCTTCATCAAGCCATGAAGGCTGTGGAAGTGCCTGCAATAGGCAATGATCAGA 1095
QY 1094 TTATGGTTCCCGTTCTGAAATGACATGAATGATCAAGCATTCATCAGAGATGT 1153
Db 1096 TTATGGTTCCCGTTCTGAAATGACATGAATGATCAAGCATTCATCAGAGATGT 1155
QY 1154 ATGTATCTCCCTTTCAGAGCCATCTCGATGAACCAAGGAGTATGAGAAATCC 1213
Db 1156 ATGTATCTCCCTTTCAGAGCCATCTCGATGAACCAAGGAGTATGAGAAATCC 1215
QY 1214 TGTGGCTCTTTTGGAGGAACTAAACATTTCCCTTGATGCTCAAGCTATGATCAGAA 1273
Db 1216 TGTGGCTCTTTTGGAGGAACTAAACATTTCCCTTGATGCTCAAGCTATGATCAGAA 1275
QY 1274 GACTTTAATATATATTTCTATCTATAGCTTAATAGGAAAGTTCTTCACAGGATT 1333
Db 1276 GACTTTAATATATATTTCTATCTATAGCTTAATAGGAAAGTTCTTCACAGGATT 1335
QY 1334 ACAGTGTAGCTACCTACATGCTGAAATAATAGCTTTAAATCATTTTATATATAACT 1393
Db 1336 ACAGTGTAGCTACCTACATGCTGAAATAATAGCTTTAAATCATTTTATATAACT 1395
QY 1394 CTGTATATAGAGATAGTCCATTTTAAATGTTTTTCCCAACCAATAAACCCCTAT 1453
Db 1396 CTGTATATAGAGATAGTCCATTTTAAATGTTTTTCCCAACCAATAAACCCCTAT 1455
QY 1454 ACAAGTTCTTCTAGTAACAATACATGAGAAAGATGCTATGTAGCTGAAATAAATGNC 1513
Db 1456 ACAAGTTCTTCTAGTAACAATACATGAGAAAGATGCTATGTAGCTGAAATAAATGNC 1515
QY 1514 GTC 1516
Db 1516 GTC 1518

```

```

RESULT 5
AAH72950
ID AAH72950 standard; cDNA; 2334 BP.
XX
AC
XX
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 4224.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
EN WO200142467-A2.
XX
XX 14-JUN-2001.
XX
PD 08-DEC-2000; 2000WO-US033312.
PF

```

```

XX 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX MPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
XX Claim 1; Page 925; 1051pp; English.
XX
XX The invention relates to novel genes (AAH6727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful; to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
XX Sequence 2334 BP; 774 A; 570 C; 510 G; 480 T; 0 U; 0 Other;
SQ

```

Query Match 98.7%; Score 1495.6; DB 4; Length 2334;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1498; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 14 CATAGCTCAGCAGCATGTACTTCAACAGAGAGGAGGAGCAATTAATACTTCTGTTGCTAGG 73
Db 16 CTGGGTGAGCAGCATGTACTTCAACAGAGAGGAGGAGCAATTAATACTTCTGTTGCTAGG 75
QY 74 TGTGGCTTCTTTAAATCTTATAAATCAGAGCCCAAGTCTCCATGCGCAGTGTGAAA 133
Db 76 TGTGGCTTCTTTAAATCTTATAAATCAGAGCCCAAGTCTCCATGCGCAGTGTGAAA 135
QY 134 TCTTCAGAGAGAGATTTCTCTTTAGTTCTTTGCAAGAGGAGTAGAGATAAAGACACTTTT 193
Db 136 TCTTCAGAGAGAGATTTCTCTTTAGTTCTTTGCAAGAGGAGTAGAGATAAAGACACTTTT 195
QY 194 CAAAAATGCAATGTATCAGAAATTCCTCAAGCAGGCTGTTTATTGAAATGAAGAGC 253
Db 196 CAAAAATGCAATGTATCAGAAATTCCTCAAGCAGGCTGTTTATTGAAATGAAGAGC 255
QY 254 AGGAATATGTTCAAACTGTGAAGTCAATCCAAAGGTGTTCCCGGATCAGCGGTGAGCCCT 313
Db 256 AGGAATATGTTCAAACTGTGAAGTCAATCCAAAGGTGTTCCCGGATCAGCGGTGAGCCCT 315
QY 314 ATCTTACCTTCAATCCATCTCGGATGTGCTGCTTGCATTAAGGCCATAATGGTTAAAG 373
Db 316 ATCTTACCTTCAATCCATCTCGGATGTGCTGCTTGCATTAAGGCCATAATGGTTAAAG 375
QY 374 GTGTGATGAAGCAACCATCATTTCTAACTTAAGCGGAAACATGCACAGCTCAAC 433
Db 376 GTGTGATGAAGCAACCATCATTTCTAACTTAAGCGGAAACATGCACAGCTCAAC 435
QY 434 AGATCAAAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTTGGATGAAACACTTGAAGAAG 493
Db 436 AGATCAAAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTTGGATGAAACACTTGAAGAAG 495
QY 494 CCCTTACAGGTCACTTGAAGAGGTGTTTGTAGTCTCTGCTAAAGAACTCAGCGCAATTG 553
Db 496 CCCTTACAGGTCACTTGAAGAGGTGTTTGTAGTCTCTGCTAAAGAACTCAGCGCAATTG 555
QY 554 ATGCTGATGAACCTCGTGTGCTGCAATGAAGGGCTTTGGAACTGATGAAGATCTCTAATTG 613
Db 556 ATGCTGATGAACCTCGTGTGCTGCAATGAAGGGCTTTGGAACTGATGAAGATCTCTAATTG 615

```



QY 614 AGATTTGGGATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGGCTTACAGAGAGG 673  
 DB 616 AGATTTGGGATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGGCTTACAGAGAGG 675  
 QY 674 AACTGAAGAGAGATCTGGCCAAAGACATCACTCAGACACATCTGGAGATTTTCGGAAAG 733  
 DB 676 AACTGAAGAGAGATCTGGCCAAAGACATCACTCAGACACATCTGGAGATTTTCGGAAAG 735  
 QY 734 CTTTGGCTTCTCTTCTTAAGGCTGACCGATCTGAGGACTTTGGTGTGAATCAGAGACTGG 793  
 DB 736 CTTTGGCTTCTCTTCTTAAGGCTGACCGATCTGAGGACTTTGGTGTGAATCAGAGACTGG 795  
 QY 794 CTGATTCAGATGCCAGGCGCTTTGATGAAGCAGAGAGAAAGGAGAAAGGACACAGCTAA 853  
 DB 796 CTGATTCAGATGCCAGGCGCTTTGATGAAGCAGAGAGAAAGGAGAAAGGACACAGCTAA 855  
 QY 854 ACGTGTTCATACCACTCTTACCAAGCAGACATCACTCAGAGAGTGTTCAGAG 913  
 DB 856 ACGTGTTCATACCACTCTTACCAAGCAGACATCACTCAGAGAGTGTTCAGAG 915  
 QY 914 AATACCAAGTACAGTAAAGTACATGACATCACTCAAGGCTTCTGAGCTGAGTGTGAAGGTTG 973  
 DB 916 AATACCAAGTACAGTAAAGTACATGACATCACTCAAGGCTTCTGAGCTGAGTGTGAAGGTTG 975  
 QY 974 ACATTCAGAAATGCTCTCAGCTATCTGTAAGTGGCCACCAAGCAACACAGCTTCTTTG 1033  
 DB 976 ACATTCAGAAATGCTCTCAGCTATCTGTAAGTGGCCACCAAGCAACACAGCTTCTTTG 1035  
 QY 1034 CAGAGAGCTTCATCAAGCCATGAAGGTTGGAACCTCGCAATGAAGGCTTGTATCAGGA 1093  
 DB 1036 CAGAGAGCTTCATCAAGCCATGAAGGTTGGAACCTCGCAATGAAGGCTTGTATCAGGA 1095  
 QY 1094 TTATGGTTCCTGGTCTGAAATGATGATATCAAGGCTTCTTATCAGAGATGT 1153  
 DB 1096 TTATGGTTCCTGGTCTGAAATGATGATATCAAGGCTTCTTATCAGAGATGT 1155  
 QY 1154 ATGGTATCTCCTTTTCCCAAGCCATCTGATGAAGCAACCAAGGAGATGATCAGAAATCC 1213  
 DB 1156 ATGGTATCTCCTTTTCCCAAGCCATCTGATGAAGCAACCAAGGAGATGATCAGAAATCC 1215  
 QY 1214 TGGTGGCTTCTTGTGGAGAACTAAACATTCCTTGTATGCTCAAGCTATGATCAGAA 1273  
 DB 1216 TGGTGGCTTCTTGTGGAGAACTAAACATTCCTTGTATGCTCAAGCTATGATCAGAA 1275  
 QY 1274 GACTTAAATATATATTTTCACTTAACTTAAATAGGAAAGTCTTCAACAGGAT 1333  
 DB 1276 GACTTAAATATATATTTTCACTTAACTTAAATAGGAAAGTCTTCAACAGGAT 1335  
 QY 1334 ACAGTGTAGCTACCTACATGCTGAAATAATATAGCCTTTTAAATCATTTTATATTAAT 1393  
 DB 1336 ACAGTGTAGCTACCTACATGCTGAAATAATATAGCCTTTTAAATCATTTTATATTAAT 1395  
 QY 1394 CTGTATATATAGATAGTCCATTTTAAATAATGTTTCCCAACCAATCAACCTTAT 1453  
 DB 1396 CTGTATATATAGATAGTCCATTTTAAATAATGTTTCCCAACCAATCAACCTTAT 1455  
 QY 1454 ACAAGTGTCTTCTAGTAAACATACATGAGAAAGATGCTATGATGCTGAAATAATATGNC 1513  
 DB 1456 ACAAGTGTCTTCTAGTAAACATACATGAGAAAGATGCTATGATGCTGAAATAATATGAC 1515  
 QY 1514 GTC 1516  
 DB 1516 GTC 1518

RESULT 6  
 ABK09784  
 ID ABK09784 standard; cDNA; 1399 BP.  
 XX  
 AC ABK09784;  
 XX  
 DT 14-MAR-2002 (first entry)

XX Human ovarian tumour protein encoding cDNA #317.  
 DE  
 XX Human; ovarian tumour protein; cancer; cytostatic; immunostimulant; ss;  
 KW gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.  
 XX Homo sapiens.  
 OS  
 XX WO200190154-A2.  
 PN  
 XX 29-NOV-2001.  
 PD  
 XX 23-MAY-2001; 2001WO-US016895.  
 PF  
 XX 24-MAY-2000; 2000US-0207107P.  
 PR  
 XX 13-JUN-2000; 2000US-0211457P.  
 PR  
 XX 21-JUN-2000; 2000US-0213673P.  
 PR  
 XX 03-AUG-2000; 2000US-0223288P.  
 PR  
 XX 01-MAR-2001; 2001US-0272790P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 XX Xu J, Mitcham JL, Harlocker SL, Dillon DC, Secretist H, Lodes MJ;  
 PI Algate PA, Fling SP, Mannion J, Benson DR, Carter D;  
 PI  
 XX WPI; 2002-097641/13.  
 DR  
 XX New isolated polynucleotide encoding polypeptide comprising portion of  
 PT ovarian tumor protein, useful for detection, diagnosis and therapy of  
 PT human ovarian cancer.  
 XX  
 XX Claim 1; Page 261-262; 285pp; English.  
 PS  
 XX The invention relates to an isolated polynucleotide encoding a  
 CC polypeptide comprising a portion of an ovarian tumour protein. The  
 CC sequences of the invention are useful for stimulating an immune response  
 CC and for treating ovarian cancer in a patient. An antigen presenting cell  
 CC that expresses the sequences is useful for treating ovarian cancer by  
 CC incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells  
 CC can then be proliferated and administered to the patient to inhibit the  
 CC development of cancer. The DNA sequences are useful as probes or primers  
 CC for nucleic acid hybridisation, to direct expression of a polypeptide in  
 CC appropriate host cells. Detecting the presence of a cancer in a patient  
 CC involves obtaining a biological sample from the patient, contacting the  
 CC biological sample with an agent that binds to the protein, detecting the  
 CC amount of protein that binds to the agent, comparing the amount of  
 CC protein to a predetermined cut-off value and determining the presence of  
 CC cancer. Sequences ABK09464-ABK09802 represent PCR primers and cDNA  
 CC molecules encoding ovarian tumour proteins of the invention  
 XX  
 SQ Sequence 1399 BP; 464 A; 273 C; 299 G; 363 T; 0 U; 0 Other;  
 Query Match 91.6%; Score 1389.4; DB 6; Length 1399;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 125 AGTGTGAATCTTCAGAGAAATTTCTCTTTAGTTCTTTGCAAGAGCTAGAGATAAG 184  
 DB 1 AGTGTGAATCTTCAGAGAAATTTCTCTTTAGTTCTTTGCAAGAGCTAGAGATAAG 60  
 QY 185 ACATCTTTTCAAAAATGGCAATGGTATCAGAAATCTCTCAAGCAGGCTGTTTATTGAAA 244  
 DB 61 ACATCTTTTCAAAAATGGCAATGGTATCAGAAATCTCTCAAGCAGGCTGTTTATTGAAA 120  
 QY 245 ATGAAGAGCAGGAATATGTTCAAACCTGTGAAGTCAATCCAAAGTGTCCCGATCAGCG 304  
 DB 121 ATGAAGAGCAGGAATATGTTCAAACCTGTGAAGTCAATCCAAAGTGTCCCGATCAGCG 180  
 QY 305 TGAGCCCTATCTTACCTTCAATCCATCTCGGATGTGCTGCTTGCATTAAGCCATTA 364  
 DB 181 TGAGCCCTATCTTACCTTCAATCCATCTCGGATGTGCTGCTTGCATTAAGCCATTA 240  
 QY 365 TGGTTAAAGGTGTGGATGAAGCAACCATCATTCATGACATTTCTTAAGCGGAAACATGAC 424



Db	1321	AAACCCATATACAAAGTGTCTTAGTAACTACATACAGAAAGATGCTCTATGTAGCTGAAA	1380
Qy	1505	TAAATGNCGTG 1516	
Db	1381	TAAATGACGTC 1392	
RESULT 7			
AAAS7357	standard; DNA; 1399 BP.		
AC	AAAS7357;		
XX			
DT	26-JUL-1999 (first entry)		
XX			
DE	Human p-40/annexin I DNA.		
XX			
XX	P-40; annexin I; human; MDR; multidrug resistance; antisense; prevention;		
KW	cancer treatment; detection; pathogen; therapeutic agent;		
KW	antifungal agent; plant; resistance; antitumour; antifungal; ss.		
XX			
OS	Homc sapiens.		
XX			
XX	W09921980-A1.		
XX			
PD	06-MAY-1999.		
XX			
PF	26-OCT-1998; 98WO-CR000992.		
XX			
XX	24-OCT-1997; 97CA-02219299.		
XX	(UYMC-) UNIV MCGILL.		
PA			
PI	Georges E, Wang Y;		
XX			
DR	WPI; 1999-337419/28.		
DR	P-PSDB; AAY08412.		
XX			
PT	Modulating or assessing multidrug resistance related to annexin proteins.		
PS	Claim 5; Fig 1B-C; 63pp; English.		
XX			
CC	This invention describes a novel human annexin family member, P-40 (also		
CC	known as annexin I) which is a member of the MDR (multidrug resistance)		
CC	gene family, for assessing or modulating MDR in a cell. Antisense P-40		
CC	sequences are used to prevent MDR in animals, particularly in conjunction		
CC	with cancer treatment. Detecting levels of the P-40 nucleic acid, or		
CC	related RNA, is used to detect cancer (or pathogens) with MDR, or		
CC	susceptibility. P-40 nucleic acid can also be used as a target for		
CC	identifying therapeutic agents, e.g. antifungal agents, and increasing		
CC	the nucleic acid expression in plants may be used to develop specific		
CC	resistance. The products of the invention have antitumour and antifungal		
CC	activity		
XX			
XX			
XX	Sequence 1399 BP; 464 A; 273 C; 298 G; 364 T; 0 U; 0 Other;		
XX			
Qy	Query Match	91.5%; Score 1387.8; DB 2; Length 1399;	
Db	Best Local Similarity	99.8%; Pred. No. 0;	
	Matches 1399; Conservative	0; Mismatches 3; Indels 0; Gaps 0;	
Qy	125	AGTGTGAATCTTCAGACGAGATTTCTCTTTAGTTCTTTGCAAGAAGTACAGATAAAG 184	
Db	1	AGTGTGAATCTTCAGACGAGATTTCTCTTTAGTTCTTTGCAAGAAGTACAGATAAAG 60	
Qy	185	ACACTTTTTCAAAAATGCGAATGGTATCAGAATTCCTCAAGCAGGCTGGTTATTGAAA 244	
Db	61	ACACTTTTTCAAAAATGCGAATGGTATCAGAATTCCTCAAGCAGGCTGGTTATTGAAA 120	
Qy	245	ATGACGACGCGGATATGTTTCCAACTGTGAAGTCATCCAAAGTGGTCCCGGATCAGCGG 304	
Db	121	ATGACGACGCGGATATGTTTCCAACTGTGAAGTCATCCAAAGTGGTCCCGGATCAGCGG 180	
Qy	305	TGAGCCCCCTATCTACCTACCTTCGAATCCATCCCTCGGATGTCGCTTGCATAGGCCATTA 364	

181	DB	TTGAGCCCTTATCTCACTTCAATTCATCTCGGATGTCCTGCTTGTGATAGGCGCATAA	240
365	QY	TGGTTAAAGGCTGGATGAAGCAACCATCATTTGACATTTCTAATGACGAAACAATGAC	424
241	DB	TGGTTAAAGGCTGGATGAAGCAACCATCATTTGACATTTCTAATGACGAAACAATGAC	300
425	QY	AGCGTCAACAGATCAAGACAGCATATCTCCAGAAACAGGAAGCCCTTGGATGAACAAC	484
301	DB	AGCGTCAACAGATCAAGACAGCATATCTCCAGAAACAGGAAGCCCTTGGATGAACAAC	360
485	QY	TGAAGAAAGCCCTTACAGCTCACCTTGAGAGGTTGTTTTAGCTCTGCTTAAATACTCCAG	544
361	DB	TTAAGAAAGCCCTTACAGCTCACCTTGAGAGGTTGTTTTAGCTCTGCTTAAATACTCCAG	420
545	QY	CGCAATTTGATGCTGATGAACCTTGTGCTGCCATGAAGGCGCTTGGAACTGATGAAGATA	604
421	DB	CGCAATTTGATGCTGATGAACCTTGTGCTGCCATGAAGGCGCTTGGAACTGATGAAGATA	480
605	QY	CTCTAAATTGAGATTTTGGCATCAAGAACTAAACAGAAATCAGAGCATTAACAGGGTCT	664
481	DB	CTCTAAATTGAGATTTTGGCATCAAGAACTAAACAGAAATCAGAGCATTAACAGGGTCT	540
665	QY	ACAGACAGCAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAGATT	724
541	DB	ACAGACAGCAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAGATT	600
725	QY	TTCGGAACGCTTTGCTTTCTTTGCTTAAAGGTGAACCGATCTGAGGACTTTTGGTGTGAATG	784
601	DB	TTCGGAACGCTTTGCTTTCTTTGCTTAAAGGTGAACCGATCTGAGGACTTTTGGTGTGAATG	660
785	QY	AGACTTTGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGAGAAAGAGAAAGGGA	844
661	DB	AGACTTTGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGAGAAAGAGAAAGGGA	720
845	QY	CAGACGTAAACGCTGTTCAATACCATCTTACCAACAGAGCTATCCAACTTCGCAGAG	904
721	DB	CAGACGTAAACGCTGTTCAATACCATCTTACCAACAGAGCTATCCAACTTCGCAGAG	780
905	QY	TGTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCCTGGACCTGGACT	964
781	DB	TGTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCCTGGACCTGGACT	840
965	QY	TGAAGGTGCATTTGAGAAATGCCTCAACAGCTATCGTGAAGTCGCGCCACAAGCAACCCAG	1024
841	DB	TGAAGGTGCATTTGAGAAATGCCTCAACAGCTATCGTGAAGTCGCGCCACAAGCAACCCAG	900
1025	QY	CTTTCTTTTGAGAGAGCTTTCATCAAGCCCATGAAGGTTTGGAACTCGCCATAAGGCAT	1084
901	DB	CTTTCTTTTGAGAGAGCTTTCATCAAGCCCATGAAGGTTTGGAACTCGCCATAAGGCAT	960
1085	QY	TGATCAGGATTAATGTTTCCGTTCTGAAATTCACATGAATGATATCAAGCACTTCTATC	1144
961	DB	TGATCAGGATTAATGTTTCCGTTCTGAAATTCACATGAATGATATCAAGCACTTCTATC	1020
1145	QY	AGAAGATGATATGSTATCTCCCTTTGCGAAGCCATCTTGGATGAACCAAGAGAGATATG	1204
1021	DB	AGAAGATGATATGSTATCTCCCTTTGCGAAGCCATCTTGGATGAACCAAGAGAGATATG	1080
1205	QY	AGAAATCTCGTGGCTCTTTGTGGAGGAACATAACATTTCCCTTGATGTCTCAAGCTA	1264
1081	DB	AGAAATCTCGTGGCTCTTTGTGGAGGAACATAACATTTCCCTTGATGTCTCAAGCTA	1140
1265	QY	TGATCAGAAGACTTTAAATTATATATTTTCATCTTAAAGCTTAAATAGGAAAGTTCTTC	1324
1141	DB	TGATCAGAAGACTTTAAATTATATATTTTCATCTTAAAGCTTAAATAGGAAAGTTCTTC	1200
1325	QY	AACAGGATTAAGTGTAGCTACCTACATGCTGAATAATATAGCTTTTAATCATTTTTAT	1384
1201	DB	AACAGGATTAAGTGTAGCTACCTACATGCTGAATAATATAGCTTTTAATCATTTTTAT	1260
1385	QY	ATTATAACTCTGTATAATAGAGATAAGTCCATTTTTTAAAAATGTTTTTCCCCCAACCAT	1444

Db	1261	ATTATAACTCTGTATTAATATAGAGATAAGTCCATTTTAAAAATGTTTCCCCCAACCATTA	1322
Qy	1445	AAACCCCTATACAAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTTCTATGTAGCTGAAAA	1504
Db	1321	AAACCCCTATACAAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTTCTATGTAGCTGAAAA	1380
Qy	1505	TAAATGTCGTC	1516
Db	1381	TAAATGACGTC	1392
RESULT 8			
ABL66809			
ID	ABL66809 standard; DNA; 1399 BP.		
XX	ABL66809;		
XX	15-MAY-2002 (first entry)		
XX	Lung cancer related gene sequence SEQ ID NO:5146.		
XX	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;		
KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;		
KW	cytostatic; Gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;		
KW	gene; ds.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	WO200194629-A2.		
PN	13-DEC-2001.		
XX	30-MAY-2001; 2001WO-US010838.		
XX	05-JUN-2000; 2000US-0209473P.		
PR	05-JUN-2000; 2000US-0209531P.		
PR	18-SEP-2000; 2000US-0233133P.		
PR	18-SEP-2000; 2000US-0233617P.		
PR	20-SEP-2000; 2000US-0234009P.		
PR	20-SEP-2000; 2000US-0234034P.		
PR	20-SEP-2000; 2000US-0234052P.		
PR	22-SEP-2000; 2000US-0234509P.		
PR	22-SEP-2000; 2000US-0234567P.		
PR	25-SEP-2000; 2000US-0234923P.		
PR	25-SEP-2000; 2000US-0234924P.		
PR	25-SEP-2000; 2000US-0235077P.		
PR	25-SEP-2000; 2000US-0235082P.		
PR	25-SEP-2000; 2000US-0235134P.		
PR	25-SEP-2000; 2000US-0235280P.		
PR	26-SEP-2000; 2000US-0235637P.		
PR	26-SEP-2000; 2000US-0235638P.		
PR	27-SEP-2000; 2000US-0235711P.		
PR	27-SEP-2000; 2000US-0235720P.		
PR	27-SEP-2000; 2000US-0235840P.		
PR	27-SEP-2000; 2000US-0235863P.		
PR	28-SEP-2000; 2000US-0236028P.		
PR	28-SEP-2000; 2000US-0236032P.		
PR	28-SEP-2000; 2000US-0236033P.		
PR	28-SEP-2000; 2000US-0236034P.		
PR	28-SEP-2000; 2000US-0236109P.		
PR	28-SEP-2000; 2000US-0236111P.		
PR	29-SEP-2000; 2000US-0236842P.		
PR	29-SEP-2000; 2000US-0236891P.		
PR	29-SEP-2000; 2000US-0237172P.		
PR	02-OCT-2000; 2000US-0237173P.		
PR	02-OCT-2000; 2000US-0237278P.		
PR	02-OCT-2000; 2000US-0237294P.		
PR	02-OCT-2000; 2000US-0237295P.		
PR	02-OCT-2000; 2000US-0237316P.		
PR	03-OCT-2000; 2000US-0237425P.		
PR	03-OCT-2000; 2000US-0237558P.		
PR	03-OCT-2000; 2000US-0237604P.		
PR	03-OCT-2000; 2000US-0237606P.		

PR 03-OCT-2000; 2000US-0237608P.  
PR 01-NOV-2000; 2000US-0244867P.  
PR 01-NOV-2000; 2000US-0245084P.  
XX (AVAL-) AVALON PHARM.  
XX Young PE, Augustus M., Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX WPI; 2002-188264/24.  
XX  
XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
PT agent to be tested for anti-neoplastic activity, and determining a change  
PT in expression of a gene of a signature gene set.  
XX  
XX Claim 1; SEQ ID NO 5146; 44pp; English.  
XX  
XX The present invention describes a method (M1) for screening for an anti-  
CC neoplastic agent. The method involves exposing cells to a chemical agent  
CC to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 847 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cycostatic  
CC activity and can be used in gene therapy. M1 can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of M1, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. M1 can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms's  
XX tumour  
XX  
XX Sequence 1399 BP; 464 A; 273 C; 298 G; 364 T; 0 U; 0 Other;  
XX  
XX Query Match 91.5%; Score 1387.8; DB 6; Length 1399;  
XX Best Local Similarity 99.8%; Pred. No. 0;  
XX Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
XX  
XX 125 AGTGTGAATCTTCAGAGAAGAAATTCCTTTAGTCTTTTCAAGAGAGGTAGAGATAAAG 184  
XX 1 AGTGTGAATCTTCAGAGAAGAAATTCCTTTAGTCTTTTCAAGAGAGGTAGAGATAAAG 60  
XX  
XX 185 ACACCTTTTCAAAAATGGCAATGTATCAGAAATTCCTCAGCAGGCTGGTTATTGAAA 244  
XX  
XX 61 ACACCTTTTCAAAAATGGCAATGTATCAGAAATTCCTCAGCAGGCTGGTTATTGAAA 120  
XX  
XX 245 ATGAGAGCAGGATATCTTCAAACTGTGAAGTCAATCCAAAGGTGGTCCCGGATCAGCGG 304  
XX  
XX 121 ATGAGAGCAGGATATCTTCAAACTGTGAAGTCAATCCAAAGGTGGTCCCGGATCAGCGG 180  
XX  
XX 305 TGAGCCCTATCTTACCTTCAATCATCTCCGATGTGCTGCTTGCATAGGCCATAA 364  
XX  
XX 181 TGAGCCCTATCTTACCTTCAATCATCTCCGATGTGCTGCTTGCATAGGCCATAA 240  
XX  
XX 365 TGGTTAAAGGTGGATGAGCAACCATCATCTTACATTTCTAAGCGAACAATGCAC 424  
XX  
XX 241 TGGTTAAAGGTGGATGAGCAACCATCATCTTACATTTCTAAGCGAACAATGCAC 300  
XX  
XX 425 AGCGTCAACAGATCAAGAGCAGCATATCTCCAGGAACAGGAAGCCCTGGATGAACAC 484  
XX  
XX 301 AGCGTCAACAGATCAAGAGCAGCATATCTCCAGGAACAGGAAGCCCTGGATGAACAC 360  
XX  
XX 485 TGAAGAGAGCCCTTACAGGTCACTTGGAGGTGGTTTGTAGCTTGTCTGCTAAAACTCCAG 544  
XX  
XX 361 TTAAGAGAGCCCTTACAGGTCACTTGGAGGTGGTTTGTAGCTTGTCTGCTAAAACTCCAG 420  
XX  
XX 545 CGCAATTTGATGCTGATGAACCTTCGTGCTCCATGAAGGGCTTGGAACTGATGAAGATA 604  
XX  
XX 421 CGCAATTTGATGCTGATGAACCTTCGTGCTCCATGAAGGGCTTGGAACTGATGAAGATA 480  
XX

QY 605 CTCTAATTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAAACAGGCTCT 664  
DB 481 CTCTAATTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAAACAGGCTCT 540  
QY 665 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATACCTCAGACACATCTTGGAGATT 724  
DB 541 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATACCTCAGACACATCTTGGAGATT 600  
QY 725 TTCCGAAAGCTTTGCTTTCTTCTTAAAGGTGACCGATCTGAGGACTTTGGTGTGAATG 784  
DB 601 TTCCGAAAGCTTTGCTTTCTTCTTAAAGGTGACCGATCTGAGGACTTTGGTGTGAATG 660  
QY 785 AAGACTTGGCTGATTCAGATGCCAGGCTTTGATGAAGCAGAGAGAAAGAGAAAGGGA 844  
DB 661 AAGACTTGGCTGATTCAGATGCCAGGCTTTGATGAAGCAGAGAGAAAGAGAAAGGGA 720  
QY 845 CAGAGCTAAAGCTGCTCAATACCATCTTACCACCAAGCTATCCACAACTTCGCGAGAG 904  
DB 721 CAGAGCTAAAGCTGCTCAATACCATCTTACCACCAAGCTATCCACAACTTCGCGAGAG 780  
QY 905 TGTTCAGAAATACACCAAGTACAGTAAAGCATGATGAACAAAGTTCTGGACCTGGAGT 964  
DB 781 TGTTCAGAAATACACCAAGTACAGTAAAGCATGATGAACAAAGTTCTGGACCTGGAGT 840  
QY 965 TGAAGGTGACATTCAGAAATGCTCAGCTATCGTGAAGTGGCCACCAAGCAAAACAG 1024  
DB 841 TGAAGGTGACATTCAGAAATGCTCAGCTATCGTGAAGTGGCCACCAAGCAAAACAG 900  
QY 1025 CTTCCTTTTCAGAGAGCTTCATCAAGCCATGAAAGGTGTGGAACTCGCCATGAAGCAT 1084  
DB 901 CTTCCTTTTCAGAGAGCTTCATCAAGCCATGAAAGGTGTGGAACTCGCCATGAAGCAT 960  
QY 1085 TGATCAGATTTATGGTTTCCGTTCTGAAATTCATGATGATCAAAAGCATTTCTATC 1144  
DB 961 TGATCAGATTTATGGTTTCCGTTCTGAAATTCATGATGATCAAAAGCATTTCTATC 1020  
QY 1145 AGAAGATGTATGGTATCTCCCTTTGCCAAGCCATCTCGATGAAACCAAGAGAGATG 1204  
DB 1021 AGAAGATGTATGGTATCTCCCTTTGCCAAGCCATCTCGATGAAACCAAGAGAGATG 1080  
QY 1205 AGAAATCTCGTGGCTCTTTGTGGAGAACTAAACATTCCTCTGATGGTCTCAAGCTA 1264  
DB 1081 AGAAATCTCGTGGCTCTTTGTGGAGAACTAAACATTCCTCTGATGGTCTCAAGCTA 1140  
QY 1265 TGATCAGAGAGCTTTAAATATATATTTTTCATCTTAAATAGCTTAAATAGGAAAGTTTCTTC 1324  
DB 1141 TGATCAGAGAGCTTTAAATATATATTTTTCATCTTAAATAGCTTAAATAGGAAAGTTTCTTC 1200  
QY 1325 AACAGATTTACAGTGTAGCTACCTACCTGCTGAAATATAGCCCTTTAAATCATTTTAT 1384  
DB 1201 AACAGATTTACAGTGTAGCTACCTACCTGCTGAAATATAGCCCTTTAAATCATTTTAT 1260  
QY 1385 ATTATACTCTCTATAATAGAGATAAGTCCATTTTTTAAAAATGTTTTTCCCCAAACCATA 1444  
DB 1261 ATTATACTCTCTATAATAGAGATAAGTCCATTTTTTAAAAATGTTTTTCCCCAAACCATA 1320  
QY 1445 AAACCCCTATCAAGTTGTTCTAGTAAACATATACATCAGAAAGATGCTATGTAGCTGAAAA 1504  
DB 1321 AAACCCCTATCAAGTTGTTCTAGTAAACATATACATCAGAAAGATGCTATGTAGCTGAAAA 1380  
QY 1505 TAAATGAGCTC 1516  
DB 1381 TAAATGAGCTC 1392  
RESULT 9  
AAS94747  
ID AAS94747 standard; DNA; 1399 BP.  
XX  
AC AAS94747;  
XX  
XX 14-FEB-2002 (first entry)  
XX

Human DNA sequence #2 expressed during foam cell differentiation.  
Human; foam cell differentiation; atherosclerosis; cerebral stroke;  
cardiovascular disorder; coronary artery disease; gene therapy; ds.  
Homo sapiens.  
W0200177389-A2.  
18-OCT-2001.  
04-APR-2001; 2001WO-US011128.  
05-APR-2000; 2000US-0195106P.  
(INCY-) INCYTE GENOMICS INC.  
Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;  
Tai J;  
WPI; 2002-010925/01.  
Composition useful for diagnosis of conditions, disorders or diseases  
associated with atherosclerosis, comprises several polynucleotides that  
are differentially expressed in foam cell development.  
Claim 1; Page 60-61; 315pp; English.  
The present invention relates to the isolation of human polynucleotide  
sequences that are differentially expressed during foam cell  
differentiation. The polynucleotide sequences of the invention or a  
composition comprising these polynucleotides are useful as a high  
throughput method for detecting altered expression of one or more  
polynucleotides in a sample. The polynucleotides can be used in the  
diagnosis of disorders associated with foam cell development such as  
atherosclerosis, cerebral stroke, and cardiovascular disorders such as  
coronary artery disease. The polynucleotide sequences can also be used as  
PCR primers and probes. The polynucleotides of the invention are also  
useful in gene therapy. AAS94746-AAS95021 represent the human  
polynucleotide sequences of the invention which are differentially  
expressed during foam cell differentiation  
Sequence 1399 BP; 464 A; 273 C; 298 G; 364 T; 0 U; 0 Other;  
Query Match 91.5%; Score 1397.8; DB 6; Length 1399;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
125 AGTGTGAATCTTCAGAGAGAAATTCCTTTAGTTCCTTTGCAAGAGGTAGAGATAAG 184  
1 AGTGTGAATCTTCAGAGAGAAATTCCTTTAGTTCCTTTGCAAGAGGTAGAGATAAG 60  
185 ACACCTTTTCAAAAATGGCAATGGTATCAGAAATCCTCAAGCAGCCCTGGTTTATTGAAA 244  
61 ACACCTTTTCAAAAATGGCAATGGTATCAGAAATCCTCAAGCAGCCCTGGTTTATTGAAA 120  
245 ATGAGAGCAGGAATATGTTTCAAACTGTAAGTCAATCAAGAGTGGTCCCGATCAGCGG 304  
121 ATGAGAGCAGGAATATGTTTCAAACTGTAAGTCAATCAAGAGTGGTCCCGATCAGCGG 180  
305 TGAGCCCTTACTCTTCAATCCATCTCGATGTGCTGCTGCTTGCATAGGCCATAA 364  
181 TGAGCCCTTACTCTTCAATCCATCTCGATGTGCTGCTGCTTGCATAGGCCATAA 240  
365 TGGTTAAAGTGGTGAAGCAACATCATTTGACATCTTAAGCAAGCAATGCAC 424  
241 TGGTTAAAGTGGTGAAGCAACATCATTTGACATCTTAAGCAAGCAATGCAC 300  
425 AGCGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGAAAGCCCTGGATGAAACAC 484  
301 AGCGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGAAAGCCCTGGATGAAACAC 360  
485 TGAAGAAAGCCCTTACAGGTCACCTTGAGGAGGTGTTTGTAGCTCTGTCTAAAACCTCCAG 544

361 TTAGAAAGCCCTTACAGTCACTTGGAGGAGTGTGTTTGTAGCTCTGTCTAAAACCTCCAG 420  
545 CGCAATTTGATGCTGATCAAACTTGTGCTGCTCCATGAAAGGCTTGGAACTGATCAAGATA 604  
421 CGCAATTTGATGCTGATCAAACTTGTGCTGCTCCATGAAAGGCTTGGAACTGATCAAGATA 480  
605 CTCTAAATTTGAGATTTTGGCATCAAGAACTAAACAAGAAATCAGAGACATTAACAGGGTCT 664  
481 CTCTAAATTTGAGATTTTGGCATCAAGAACTAAACAAGAAATCAGAGACATTAACAGGGTCT 540  
665 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATTAACCTCAGACACATCTGGAGATT 724  
541 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATTAACCTCAGACACATCTGGAGATT 600  
725 TTGCGAAGCCTTTGCTTCTCTTGTAAAGGGTGAACGATCTGAGGACTTTGGTGTGAATG 784  
601 TTGCGAAGCCTTTGCTTCTCTTGTAAAGGGTGAACGATCTGAGGACTTTGGTGTGAATG 660  
785 AAGACTTGGCTGATTCAGATGCCAGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA 844  
661 AAGACTTGGCTGATTCAGATGCCAGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA 720  
845 CAGAGTAAACCTGTTCAATACCATCTTACCACAGAGAGCTATCCACAACTTCCAGAG 904  
721 CAGAGTAAACCTGTTCAATACCATCTTACCACAGAGAGCTATCCACAACTTCCAGAG 780  
905 TGTTCAGAAATACACCAAGTACAGTACATGAACAAAGTCTTGGACCTGGAGT 964  
781 TGTTCAGAAATACACCAAGTACAGTACATGAACAAAGTCTTGGACCTGGAGT 840  
965 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCGTGAAGTGCAGCAACCAACCAAG 1024  
841 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCGTGAAGTGCAGCAACCAACCAAG 900  
1025 CTTTCTTTTCAGAGAGCTTCATCAAGCCATGAAGGTTGTGAACTCCCATTAAGGCT 1084  
901 CTTTCTTTTCAGAGAGCTTCATCAAGCCATGAAGGTTGTGAACTCCCATTAAGGCT 960  
1085 TGATCAGGATATGTTTCCCTGCTTGAATTTGACATGAATGATATCAAGCAATTTCTATC 1144  
961 TGATCAGGATATGTTTCCCTGCTTGAATTTGACATGAATGATATCAAGCAATTTCTATC 1020  
1145 AGAAGATGATGATCTCCCTTTGCCAAGCCATCTCGATGAAACCAACCAAGAGAGATG 1204  
1021 AGAAGATGATGATCTCCCTTTGCCAAGCCATCTCGATGAAACCAACCAAGAGATG 1080  
1205 AGAATATCCTGCTGCTCTTGTGAGGAACTTAACATTCCTTGTGATGCTCAAGCTA 1264  
1081 AGAATATCCTGCTGCTCTTGTGAGGAACTTAACATTCCTTGTGATGCTCAAGCTA 1140  
1265 TGATCAGAGACTTTAAATATATATATTTTTCATCTAATAGCTTAAATAGGAAATTTCTTC 1324  
1141 TGATCAGAGACTTTAAATATATATATTTTTCATCTAATAGCTTAAATAGGAAATTTCTTC 1200  
1325 AACAGGATACAGTGTAGCTACCTACATGCTGAAATATAGCTTTAAATCATTTTAT 1384  
1201 AACAGGATACAGTGTAGCTACCTACATGCTGAAATATAGCTTTAAATCATTTTAT 1260  
1385 ATTATACTCTGTATATAGAGATAAGTCCATTTTAAATATAGCTTTTCCCAACCAATA 1444  
1261 ATTATACTCTGTATATAGAGATAAGTCCATTTTAAATATAGCTTTTCCCAACCAATA 1320  
1445 AAACCTTATCAAGTGTCTCTAGTAAATATATATAGAGATAAGTCTATGTAGCTGAAA 1504  
1321 AAACCTTATCAAGTGTCTCTAGTAAATATATATAGAGATAAGTCTATGTAGCTGAAA 1380  
1505 TAAATGACGTC 1516  
1381 TAAATGACGTC 1392

ABX84511  
ID ABX84511 standard; cDNA; 1399 BP.

XX AC ABX84511;

XX DT 14-AUG-2002 (first entry)

XX DE Human cDNA differentially expressed in granulocytic cells #1082.

XX KW Human; ss; granulocytic cell; DNA chip; bacterial infection;  
viral infection; parasitic infection; protozoal infection;  
fungal infection; sterile inflammatory disease; psoriasis;  
rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
cardiac reperfusion injury; renal reperfusion injury; ARDS;  
adult respiratory distress syndrome; inflammatory bowel disease;  
Crohn's disease; ulcerative colitis; periodontal disease;  
granulocyte activation; chronic inflammation; allergy.

XX OS Homo sapiens.

XX PN WO200228999-A2.

XX PD 11-APR-2002.

XX PF 03-OCT-2001; 2001WO-US030821.

XX PR 03-OCT-2000; 2000US-0237189P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX DR WPI; 2002-43528/46.

XX PT Detecting granulocyte activation by detecting differential expression of  
genes associated with granulocyte activation, which serves as diagnostic  
markers that is useful for monitoring disease states and drug toxicity.

XX PS Claim 1; SEQ ID NO 1082; 114pp; English.

CC The invention relates to detecting (M1) granulocyte (GC) activation  
(GCA) by detecting the level of expression of gene(s) (Gs) identified by  
DNA chip analysis as given in the specification, and comparing the  
expression level to an expression level in an unactivated GC, where  
differential expression of Gs is indicative of GCA. Also included are  
modulating (M2) GA by contacting GC with an agent that alters the  
expression of at least one gene in Gs; (2) screening (M3) for an agent  
capable of modulating GCA or an inflammation (especially chronic) in a  
tissue, an allergic response in a subject, exposure of a subject to a  
pathogen or sterile inflammatory disease using the gene expression  
profile; (3) detecting (M4) an inflammation (especially chronic) in a  
tissue, an allergic response in a subject, exposure of a subject to a  
pathogen or sterile inflammatory disease, by detecting the level of  
expression in a sample of the tissue of gene(s) from Gs, where the level  
of expression of the gene is indicative of inflammation; (4) treating  
(M5) an inflammation (especially chronic) or in a tissue, an allergic  
response in a subject, exposure of a subject to a pathogen or sterile  
inflammatory disease, by contacting a tissue having inflammation with an  
agent that modulates the expression of gene(s) from Gs in the tissue. M1  
is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful  
for screening an agent capable of modulating GCA preferably in an  
inflammation in a tissue; M4 is useful for detecting an inflammation  
(especially chronic) in a tissue, an allergic response in a subject,  
exposure of a subject to a pathogen or sterile inflammatory disease (e.g.  
psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,  
cardiac reperfusion injury, renal reperfusion injury, ARDS, adult  
respiratory distress syndrome, inflammatory bowel disease, Crohn's  
disease, ulcerative colitis, periodontal disease; also bacterial  
infection, viral infection, parasitic infection, protozoal infection,  
fungal infection and M5 is useful for treating one of the above  
conditions. The present sequence represents a gene differentially  
expressed in granulocytes. Note: The sequence data for this patent did  
not form part of the printed specification, but was obtained in

CC electronic format directly from WIFO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 1399 BP; 464 A; 273 C; 298 G; 364 T; 0 U; 0 Other;

Query Match 91.5%; Score 1387.8; DB 6; Length 1399;  
Best Local Similarity 99.8%; Pred. NO. 0;  
Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 125 AGTGTGAATCTTCAGAGAGAAATTTCTCTTTAGTTCTTTTGAAGAAGCTAGAGATAAG 184

DB 1 AGTGTGAATCTTCAGAGAGAAATTTCTCTTTAGTTCTTTTGAAGAAGCTAGAGATAAG 60

QY 185 ACATCTTTTCAAAAATGGCAATGGTATCAGATTCCTCAAGCAGGCTGTTTATTGAAA 244

DB 61 ACATCTTTTCAAAAATGGCAATGGTATCAGATTCCTCAAGCAGGCTGTTTATTGAAA 120

QY 245 ATGAAGAGCAGGAATATGTTTCAAACTGTGAAGTCATCCAAAGTGTCCTCCGGATCAGCG 304

DB 121 ATGAAGAGCAGGAATATGTTTCAAACTGTGAAGTCATCCAAAGTGTCCTCCGGATCAGCG 180

QY 305 TGAGCCCTTATCTTACCTTCAATCCATCTCGGATGTCCTGCTTGCATTAAGGCCATAA 364

DB 181 TGAGCCCTTATCTTACCTTCAATCCATCTCGGATGTCCTGCTTGCATTAAGGCCATAA 240

QY 365 TGGTTAAAGTGATGATGAAGCAACCATCATTAAGCAATCTAACTAAGCGAAACAATGCAAC 424

DB 241 TGGTTAAAGTGATGATGAAGCAACCATCATTAAGCAATCTAACTAAGCGAAACAATGCAAC 300

QY 425 AGGTCACACAGATCAAAAGCAGAGATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 484

DB 301 AGGTCACACAGATCAAAAGCAGAGATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 360

QY 485 TGAAGAAGCCCTTACAGGTACCTTGAAGAGGTTGTTTGTAGCTCTGTCTGTTAAACCTCCAG 544

DB 361 TTAAGAAGCCCTTACAGGTACCTTGAAGAGGTTGTTTGTAGCTCTGTCTGTTAAACCTCCAG 420

QY 545 CGCAATTTGATGCTGATGAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604

DB 421 CGCAATTTGATGCTGATGAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY 605 CTCTAATGAGATTTGGCATCAAGAACTAAACAAGAAATCAGAGACATTAACAGGGTCT 664

DB 481 CTCTAATGAGATTTGGCATCAAGAACTAAACAAGAAATCAGAGACATTAACAGGGTCT 540

QY 665 ACAGAGAGGAACCTGAAGAGAGATCTGCCCAAGACATAACCTCAGACACATCTCGAGATT 724

DB 541 ACAGAGAGGAACCTGAAGAGAGATCTGCCCAAGACATAACCTCAGACACATCTCGAGATT 600

QY 725 TTGGAACCGTTTGTCTTCTTGTCTAAGGGTGACCGATCTGAGACTTTGGTGTGAATG 784

DB 601 TTGGAACCGTTTGTCTTCTTGTCTAAGGGTGACCGATCTGAGACTTTGGTGTGAATG 660

QY 785 AAGACTTGGCTGATTCCAGATGCCAGGCTTGTGTATGAGCAGGAGAAAGGAGAAAGGGA 844

DB 561 AAGACTTGGCTGATTCCAGATGCCAGGCTTGTGTATGAGCAGGAGAAAGGAGAAAGGGA 720

QY 845 CAGACGTAAACGTGTTTCAATACATCTTACACAGAGAGCTATCCACACTTCCAGAG 904

DB 721 CAGACGTAAACGTGTTTCAATACATCTTACACAGAGAGCTATCCACACTTCCAGAG 780

QY 905 TGTTTTCAGAAATACACCAAGTACAGTGAAGCATGAACAAAGTCTTGGACCTTGGAGT 964

DB 781 TGTTTTCAGAAATACACCAAGTACAGTGAAGCATGAACAAAGTCTTGGACCTTGGAGT 840

QY 965 TGAAGGTGACATTTGAGAAATGCTTCAGAGCTATCGTGAAGTGGCCCAAGCAACACAG 1024

DB 841 TGAAGGTGACATTTGAGAAATGCTTCAGAGCTATCGTGAAGTGGCCCAAGCAACACAG 900

QY 1025 CTCTTCTTTCAGAGAGAGCTTTCATCAAGCCATCAAGAGGTTGTTGGAACCTGCCCAAGGAT 1084

DB 901 CTCTTCTTTCAGAGAGAGCTTTCATCAAGCCATCAAGAGGTTGTTGGAACCTGCCCAAGGAT 960





661 AAGCTTGGCTGATTGATGAGTCCAGGCGCTTGTATGAGCAGGAGAAAGGAGGGA 720  
845 CAGACGTAAACGTTGTTCAATACATCCTTACACACAGAGAGCTATCCACAACTTCGCAGAG 904  
721 CAGACGTAAACGTTGTTCAATACATCCTTACACACAGAGAGCTATCCACAACTTCGCAGAG 780  
905 TGTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGT 964  
781 TGTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGT 840  
965 TGAAGGTGACATGTGAAGATGCTACAGCTATCTGAAAGTGGCGCCACAGCAACCAAG 1024  
841 TGAAGGTGACATGTGAAGATGCTACAGCTATCTGAAAGTGGCGCCACAGCAACCAAG 900  
1025 CTTTCCTTTCAGAGAACTTCATCAAGGCCATGAAGGTTGTTGAAGTCCGCCATTAAGCAT 1084  
901 CTTTCCTTTCAGAGAACTTCATCAAGGCCATGAAGGTTGTTGAAGTCCGCCATTAAGCAT 960  
1085 TGATCAGGATATGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCC 1144  
961 TGATCAGGATATGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCC 1020  
1145 AGAAGATGATGTTATCT 1204  
1021 AGAAGATGATGTTATCT 1080  
1205 AGAAATCCTGGTGGCTCTTTGTTGGAGGAACTAAACATTCCTTGATGGTCTCAAGCTA 1264  
1081 AGAAATCCTGGTGGCTCTTTGTTGGAGGAACTAAACATTCCTTGATGGTCTCAAGCTA 1140  
1265 TGATCAGAGACTTTAAT 1324  
1141 TGATCAGAGACTTTAAT 1200  
1325 AACAGGATACAGTGTAGTACCTACATGCTGAAATATAGCTTTAATATATATATATATAT 1384  
1201 AACAGGATACAGTGTAGTACCTACATGCTGAAATATAGCTTTAATATATATATATATAT 1260  
1385 ATTATACCTCTGATATAGATAGTACCTATTTTAAATATATATATATATATATATATATAT 1444  
1261 ATTATACCTCTGATATAGATAGTACCTATTTTAAATATATATATATATATATATATATAT 1320  
1445 AACCCCTATACAGTGTGTTCTAGTAAACATACATGAGAAAGATGTTCTATAGTCTGAAAA 1504  
1321 AACCCCTATACAGTGTGTTCTAGTAAACATACATGAGAAAGATGTTCTATAGTCTGAAAA 1380  
1505 TAAATGTCGTC 1516  
1381 TAAATGTCGTC 1392

RESULT 12

ID AAC77652 standard; cDNA; 1431 BP.

XX AC AAC77652;

XX AC AAC77652;

XX AC AAC77652;

XX AC AAC77652;

08-FEB-2001 (first entry)

Human cancer associated gene sequence SEQ ID NO:46.

Human; cancer associated gene; cancer antigen; detection; cancer;  
diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;  
antiidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;  
antiinflammatory; antithyroid; antiallergic; antidiabetic; cardiant;  
dermatologic; neuroprotective; thrombolytic; coagulant; neutropenic;  
vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;  
immune disorder; haematopoietic cell disorder; autoimmune disorder;  
allergic reaction; graft versus host disease; organ rejection;  
haemostatic; thrombolytic; cardiovascular disorder; infection;  
neurological disease; drug screening; ss.

XX Homo sapiens.

XX WO200055350-A1.  
PN 21-SEP-2000.  
XX 08-MAR-2000; 2000MO-US005882.  
XX 12-MAR-1999; 99US-0124270P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Ruben SM;  
PI WPI; 2000-587533/55.  
XX P-PSDB; AAB43443.  
DR Novel isolated nucleic acids comprising sequences encoding peptides  
PT useful for treating or diagnosing e.g. cancer.  
PS Claim 1; Page 644; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given in  
CC AAB43398 to AAB44239. The proteins can have activities based on the  
CC tissues and cells the genes are expressed in. Example of activities  
CC include: cytostatic; proliferative; vulnerability; immunomodulator;  
CC antidiabetic; antiaesthetic; antirheumatic; antiarthritic;  
CC antiinflammatory; antithyroid; antiallergic; antidiabetic; antiviral;  
CC dermatologic; neuroprotective; cardiant; thrombolytic; coagulant;  
CC neutropenic; vasotropic; antipsoriatic and antiangiogenic. The  
CC polynucleotides and polypeptides can be used for preventing, treating or  
CC ameliorating medical conditions and diagnosing pathological conditions.  
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
CC the present invention may be used to treat immune disorders by activating  
CC or inhibiting the proliferation, differentiation or mobilisation of  
CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
CC disorders, allergic reactions, graft versus host disease, and organ  
CC rejection, modulate haemostatic or thrombolytic activity, modulate  
CC inflammation, cancers, cardiovascular disorders, neurological disease and  
CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
CC the present invention

XX Sequence 1431 BP; 485 A; 281 C; 303 G; 362 T; 0 U; 0 Other;

Query Match 91.0%; Score 1379.4; DB 3; Length 1431;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1380; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 135 CTTTCAGAGAAATTTCTTTAGTTCTTTAGTTCTTTAGTTCTTTAGTTCTTTAGTTCTTT 194  
DB 20 CTTTCAGAGAAATTTCTTTAGTTCTTTAGTTCTTTAGTTCTTTAGTTCTTTAGTTCTTT 79  
QY 195 AAAAAATGCAATGGTATCAGAAATTCCTCAAGCAGCGCTGGTTTATTGAAAAATGAAGACA 254  
DB 80 AAAAAATGCAATGGTATCAGAAATTCCTCAAGCAGCGCTGGTTTATTGAAAAATGAAGACA 139  
QY 255 GGAATATGTTCAACTGTGAAGTATCCAAAGGTGGTCCCGATCAGCGGTAGCCCTTA 314  
DB 140 GGAATATGTTCAACTGTGAAGTATCCAAAGGTGGTCCCGATCAGCGGTAGCCCTTA 199  
QY 315 TCCTACCTTCATCCATCCTCGGATGCTGCTTGCATTAAGGCCATAATGGTTTAAAGG 374  
DB 200 TCTACCTTCATCCATCCTCGGATGCTGCTTGCATTAAGGCCATAATGGTTTAAAGG 259  
QY 375 TGTGATGAAGCAACCATCATTTGACATTTCTAACTAAGCGAAACATGACAGCGTCAACA 434  
DB 260 TGTGATGAAGCAACCATCATTTGACATTTCTAACTAAGCGAAACATGACAGCGTCAACA 319  
QY 435 GATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAAACACTGAAGAAAGC 494  
DB 320 GATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAAACACTGAAGAAAGC 379



Qy	495	CCTTACAGGTCACCTTTGAGGAGGTGCTTTTAGCTCTGCTATAAAACCTCCAGCGCAATTGA	554
Db	380	CCTTACAGGTCACCTTTGAGGAGGTGCTTTTAGCTCTGCTATAAAACCTCCAGCGCAATTGA	439
Qy	555	TGCTGATGAACCTTCGTGCTGCGCATGAAGGGCTTTGGAACTGATGAAGATACCTTAATTGA	614
Db	440	TGCTGATGAACCTTCGTGCTGCGCATGAAGGGCTTTGGAACTGATGAAGATACCTTAATTGA	499
Qy	615	GAATTTGGCATCAAGAACTTAACAAAGAAATCGAGACATTAACAGGGTCTACAGAGAGA	674
Db	500	GAATTTGGCATCAAGAACTTAACAAAGAAATCAAGACATTAACAGGGTCTACAGAGAGA	559
Qy	675	ACTGAAGAGAGATCTGGCCAAAGACATAACCTTCAGACACATCTGGAGATTTTCGGAACGC	734
Db	560	ACTGAAGAGAGATCTGGCCAAAGACATAACCTTCAGACACATCTGGAGATTTTCGGAACGC	619
Qy	735	TTTGCTTTCTTCGTAAAGGGTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTGGC	794
Db	620	TTTGCTTTCTTCGTAAAGGGTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTGGC	679
Qy	795	TGATTTCAGATGCGCAGGGCCCTTGATGAAGCAGAGAAAGGAGAAAGGGGACAGACGTAAA	854
Db	680	TGATTTCAGATGCGCAGGGCCCTTGATGAAGCAGAGAAAGGAGAAAGGGGACAGACGTAAA	739
Qy	855	CGTGTTCAATACCATCCTTACCAGAAAGCTATCCACACTTCGCGAGAGTGTTCAGAA	914
Db	740	CGTGTTCAATACCATCCTTACCAGAAAGCTATCCACACTTCGCGAGAGTGTTCAGAA	799
Qy	915	ATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTCGGACTCGAGTTGAAAGGTGA	974
Db	800	ATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTCGGACTCGAGTTGAAAGGTGA	859
Qy	975	CATTGAGAAATGCTCTCAGACTGCTGGAAGTCGGCCACAGCAAAACCGACTTCTTTGC	1034
Db	860	CATTGAGAAATGCTCTCAGACTGCTGGAAGTCGGCCACAGCAAAACCGACTTCTTTGC	919
Qy	1035	AGAGAAGGTTTCATCAAGCCATGAAAGGTGTTGGAATCTCGCCATAAGGCATGATCAGGAT	1094
Db	920	AGAGAAGGTTTCATCAAGCCATGAAAGGTGTTGGAATCTCGCCATAAGGCATGATCAGGAT	979
Qy	1095	TATGGTTTCGGTTCTGAAATTGACATGAATATCAAAAGCATTCCTATCAGAAAGATGTA	1154
Db	980	TATGGTTTCGGTTCTGAAATTGACATGAATATCAAAAGCATTCCTATCAGAAAGATGTA	1039
Qy	1155	TGGTATCTCCCTTTGCCAAGCCATCCTGGATGAAACCAAAGGAGATGATGAGAAAATCCT	1214
Db	1040	TGGTATCTCCCTTTGCCAAGCCATCCTGGATGAAACCAAAGGAGATGATGAGAAAATCCT	1099
Qy	1215	GGTGGCTCTTTGTGGGAGAACTAAACATTCCTTGATGGTCTCAAGCTATGATCAGAAG	1274
Db	1100	GGTGGCTCTTTGTGGGAGAACTAAACATTCCTTGATGGTCTCAAGCTATGATCAGAAG	1159
Qy	1275	ACTTTAAATTATATATTTTCATCTCTATAAGCTATAATAGGAAGTTCTTCAACAGGATTA	1334
Db	1160	ACTTTAAATTATATATTTTCATCTCTATAAGCTATAATAGGAAGTTCTTCAACAGGATTA	1219
Qy	1335	CAGCTAGCTACCTACATGCTGAAAAATATAGCCCTTTAAATCATTTTTTATTTATATAACTC	1394
Db	1220	CAGCTAGCTACCTACATGCTGAAAAATATAGCCCTTTAAATCATTTTTTATTTATATAACTC	1279
Qy	1395	TCGTATAATAGAGATAAGTCCATTTTTTAAAAAATGTTTTTCCCAAAACCAATAAAACCCCTATA	1454
Db	1280	TGTATAATAGAGATAAGTCCATTTTTTAAAAAATGTTTTTCCCAAAACCAATAAAACCCCTATA	1339
Qy	1455	CAAAGTGTGTTCTAGTAACCAATACATCAGAAAGATGCTATGTAGCTGAAAAATAAAATGCG	1514
Db	1340	CAAAGTGTGTTCTAGTAACCAATACATCAGAAAGATGCTATGTAGCTGAAAAATAAAATGCG	1399
Qy	1515	TC	1516
Db	1400	TC	1401

RESULT 13	
ADD18446	
ID	ADD18446 standard; DNA; 1377 BP.
XX AC	
XX ADD18446;	
XX AC	
XX XX	
DT DT	15-JAN-2004 (first entry)
XX XX	
DE XX	Human prostate cancer diagnosis related DNA sequence SeqID18.
XX KW	prostate tissue; cancer diagnostic; cancer marker; prostate cancer; PCA;
KW KW	male cancer-related death; serum biomarker; tissue biomarker; cytostatic;
KW KW	gene therapy; prostate biopsy tissue; AMACR;
KW KW	alpha-methylacyl-coenzyme A racemase; diagnosing cancer; cell growth;
KW KW	human; ds.
XX OS	Homo sapiens.
XX XX	
FN FN	WO2003012067-A2.
XX PD	13-FEB-2003.
XX PF	02-AUG-2002; 2002WO-US024567.
XX PR	02-AUG-2001; 2001US-0309581P.
PR PR	15-NOV-2001; 2001US-0334468P.
PR PR	01-AUG-2002; 2002US-00210120.
XX PA	(UNMI ) UNIV MICHIGAN.
XX PI	Rubin MA, Chinnaiyan AM, Sreekumar A;
XX DR	WPI; 2003-278396/27.
XX PT	Characterizing prostate tissue comprises providing a prostate tissue
PT PT	sample from a subject and detecting the presence or absence of expression
PT PT	of hepin, p1m-1 or EZH2.
XX PS	Disclosure; SEQ ID NO 18; 297pp; English.
XX CC	This invention relates to a novel method of characterising prostate
CC CC	tissue in a subject and to compositions and methods for cancer
CC CC	diagnostics, including cancer markers, in particular prostate cancer.
CC CC	Prostate cancer (pCA) is a leading cause of male cancer-related death.
CC CC	Additional serum and tissue biomarkers would aid diagnosis. The invention
CC CC	may provide means of producing compounds with a cytostatic activity or
CC CC	allow the development of gene therapy. The methods of the invention
CC CC	useful for characterising prostate tissue in a subject, screening
CC CC	compounds, characterising inconclusive prostate biopsy tissue in a
CC CC	subject, detecting AMACR (alpha-methylacyl-coenzyme A racemase)
CC CC	expression in a bodily fluid, characterising tissue in a subject,
CC CC	diagnosing cancer in a subject and inhibiting the growth of cells. The
CC CC	presented sequence is a DNA sequence which is preferably utilised in the
XX SQ	Sequence 1377 BP; 455 A; 271 C; 293 G; 358 T; 0 U; 0 Other;
Query Match	90.2%; Score 1367.4; DB 9; Length 1377;
Best Local Similarity	99.9%; Pred.No. 0;
Matches 1368; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
OY 147	ATTCTCTTTAGTTCCTTTGCAAGAAGGTGATAGATAAAGACACTTTTCAAATGGCAAT 206
DB 1	ATTCTCTTTAGTTCCTTTGCAAGAAGGTGATAGATAAAGACACTTTTCAAATGGCAAT 60
OY 207	GGTATCAGAAATTCCTCAAGCAGCGCTGGTTTTATTGAAAATGAAGACGAGGAATATGTTCA 266
DB 61	GGTAICAGAAATTCCTCAAGCAGCGCTGGTTTTATTGAAAATGAAGACGAGGAATATGTTCA 120
OY 267	AACGTGGAAGTCATCCAAAGGTGGTCCCAGATCAGCGGTAGCCCCCTATACCTCTTCAA 326
DB 121	AACGTGGAAGTCATCCAAAGGTGGTCCCAGATCAGCGGTAGCCCCCTATACCTCTTCAA 180

327 TCCATCTCGGATGTCGCTGCGCTTGCATTAAGGCCATAATGGTTAAAGGTGCGATGAAGC 386  
 181 TCCATCTCGGATGTCGCTGCGCTTGCATTAAGGCCATAATGGTTAAAGGTGCGATGAAGC 240  
 387 AACCATCATTTGACATTTCACTTAAGGGAACCAATGACAGCGCTCAACAGATCAAAGCAGC 446  
 241 AACCATCATTTGACATTTCACTTAAGGGAACCAATGACAGCGCTCAACAGATCAAAGCAGC 300  
 447 ATATCTCCAGGAACAGGAAGCCCTGGATGAACACTGAAGAAAGCCCTTACAGGTCA 506  
 301 ATATCTCCAGGAACAGGAAGCCCTGGATGAACACTGAAGAAAGCCCTTACAGGTCA 360  
 507 CCTTGAGGAGTTGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 566  
 361 CCTTGAGGAGTTGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 567 TCGTGTCCATGAAGGCTTGGAACTGATGAAGATCTCTAAATGAGATTTGGCATC 626  
 421 TCGTGTCCATGAAGGCTTGGAACTGATGAAGATCTCTAAATGAGATTTGGCATC 480  
 527 AAGNACTTAACAAAGAAATCAGAGACATTAACAGGCTTACAGAGGAACTGAAGAGAGA 686  
 481 AAGNACTTAACAAAGAAATCAGAGACATTAACAGGCTTACAGAGGAACTGAAGAGAGA 540  
 687 TCTGGCCAAAGACATAACCTCAGACACATCTGGAGATTTTCGAAAGCTTTGCTTCTCT 746  
 541 TCTGGCCAAAGACATAACCTCAGACACATCTGGAGATTTTCGAAAGCTTTGCTTCTCT 600  
 747 TGCTAAGGTTGACCGATCTCAGACATTTGGTGTGAATGAAGATTTGGTGTGAATGAAG 806  
 601 TGCTAAGGTTGACCGATCTCAGACATTTGGTGTGAATGAAGATTTGGTGTGAATGAAG 660  
 807 CAGGCGCTTGTATGAAGCAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 866  
 661 CAGGCGCTTGTATGAAGCAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
 867 CATCTTACACACAGAGCTTCCACAACTTCGAGAGTGTTCAGAAATACACCAAGTA 926  
 721 CATCTTACACACAGAGCTTCCACAACTTCGAGAGTGTTCAGAAATACACCAAGTA 780  
 927 CAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTGAAGAGGTGACATTCAGAAATG 986  
 781 CAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTGAAGAGGTGACATTCAGAAATG 840  
 987 CCTCAGAGCTATCGTAAGTGGCGGACCAAGCAACCAAGCTTTCTTTCAGAGAGGTTCA 1046  
 841 CCTCAGAGCTATCGTAAGTGGCGGACCAAGCAACCAAGCTTTCTTTCAGAGAGGTTCA 900  
 1047 TCAAGCCATGAAGGTGTTGGAATCGGCATPAAGGCATTCATCAGGATTTATGGTTCCCG 1106  
 901 TCAAGCCATGAAGGTGTTGGAATCGGCATPAAGGCATTCATCAGGATTTATGGTTCCCG 960  
 1107 TTCTGAAATTCAGATGATATCAAGCATTTCTATCAAGAGATGATGGTATCTCCCT 1166  
 961 TTCTGAAATTCAGATGATATCAAGCATTTCTATCAAGAGATGATGGTATCTCCCT 1020  
 1167 TTGCCAGCCATCTGGATGAACCAAGAGGAGATGATGAGAAATCTCGTGGCTCTTTG 1226  
 1021 TTGCCAGCCATCTGGATGAACCAAGAGGAGATGATGAGAAATCTCGTGGCTCTTTG 1080  
 1227 TGGAGGAACATAACATTCCTTGTGCTCTCAAGCTATGATCAAGAGCTTTTAATTAATA 1286  
 1081 TGGAGGAACATAACATTCCTTGTGCTCTCAAGCTATGATCAAGAGCTTTTAATTAATA 1140  
 1287 TATTTTTCATCTTAAGCTTAATAGGAAGAGTTTCTTCAAGAGCTATGATCAAGAGCTTTA 1346  
 1141 TATTTTTCATCTTAAGCTTAATAGGAAGAGTTTCTTCAAGAGCTATGATCAAGAGCTTTA 1200  
 1347 CTACATGCTGAAATATAGCTTTAAATCAATTTTATATTAATTAATTAATTAATTAATAG 1406  
 1201 CTACATGCTGAAATATAGCTTTAAATCAATTTTATATTAATTAATTAATTAATTAATAG 1260  
 1407 ATAAGTCCATTTTAAAAATGTTTTCGCCCAACCAACCAACCAACCAACCAACCAACCAAC 1466

1261 ATAAAGTCCATTTTAAAAATGTTTCCCAACACCATAAACCCCTATACAAAGTTGTTCTA 1320  
 1467 GTAACATATCATGAGAAAGATGCTATGCTAGCTGAAATAAAATGNCGTC 1516  
 1321 GTAACATATCATGAGAAAGATGCTATGCTAGCTGAAATAAAATGNCGTC 1370

RESULT 14  
 AAN60555 standard; cDNA; 1376 BP.  
 XX AAN60555;  
 XX 08-AUG-1991 (first entry)  
 XX Sequence of the cDNA insert of lambda-LC which codes for human  
 DE lipocortin.  
 XX Anti-inflammatory agent; steroid mediation; arthritis therapy; allergy;  
 KW ds.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FT CDS 1..1092  
 FT /\*tag= a  
 XX W08604094-A.  
 XX 17-JUL-1986.  
 XX 10-JAN-1986; 86WO-US000027.  
 XX 10-JAN-1985; 85US-00690146.  
 PR 15-MAR-1985; 85US-00712376.  
 PR 14-AUG-1985; 85US-00765877.  
 PR 05-SEP-1985; 85US-00772892.  
 XX  
 PA (BIOJ) BIOGEN NV.  
 XX Wallner BP, Pepinsky BR, Garwin JL, Schindler DG, Huang KS;  
 XX WPI: 1986-196888/30.  
 DR P-PSDB; AAP60657.  
 XX New human lipocortin-like polypeptide(s) - are obtd. by recombinant DNA  
 PT techniques and are antiinflammatory agents without usual side effects.  
 XX Claim 2; Page 74-75 and Fig 4; 113pp; English.  
 XX A human cDNA library of Escherichia coli cells contg. human macrophage  
 CC cDNA sequences inserted into a phage cloning vector was screened using  
 CC antisense oligonucleotide DNA probes corresp. to those regions of rat  
 CC phospholipase A2 inhibitor protein having minimal nucleotide degeneracy  
 CC (AAN60561-N60564) to obtain a sequence coding for human lipocortin (LC)-  
 CC like polypeptide  
 XX  
 XX Sequence 1376 BP; 454 A; 271 C; 292 G; 359 T; 0 U; 0 Other;  
 PS  
 Query Match 90.0%; Score 1364.8; DB 1; Length 1376;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1366; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

148 TTCTCTTTAGTTCCTTTCAGAGAGGTAGAGATAAGACACTTTTCAAAAATGCAATG 207  
 1 TTCTCTTTAGTTCCTTTCAGAGAGGTAGAGATAAGACACTTTTCAAAAATGCAATG 60  
 208 GTATCAGATTTCTCAAGCAGGCTGTTTATTTGAATGAAGCAGCAATATCTTCAA 267  
 61 GTATCAGATTTCTCAAGCAGGCTGTTTATTTGAATGAAGCAGCAATATCTTCAA 120  
 268 ACTGTGAAGTTCATCCAAAGGTGTCGCCGATCAGCGGTGAGCCCTATCTACCTTCAAT 327



these plasmids, pL5/20, was used as probe to rescreen the original library. Seventy-two plaques were isolated and analysed by Southern blotting using a 30mer (lipo 16; tag d) as a probe. The 14 clones isolated were sequenced and used to determine the full sequence encoding lipocortin. Recombinant lipocortin prep. using the DNA sequence has antiinflammatory activity and can be used for the treatment of arthritic, allergic, dermatologic, opthalmic and collagen diseases. See also CC AAQ23217-40 and AAQ23294. (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 1376 BP; 454 A; 271 C; 292 G; 359 T; 0 U; 0 Other;

Query Match 90.08; Score 1364.8; DB 2; Length 1376;  
 Best Local Similarity 99.88; Pred. No. 0;  
 Matches 1366; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

148 TTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAGACACTTTTCAAAAATGGCAATG 207  
 1 TTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAGACACTTTTCAAAAATGGCAATG 60

208 GTATCAGAAATCTCTCAGCAGGCTGTTTATTGAAATGAAGCAGGAGGAATATGTTCAA 267  
 61 GTATCAGAAATCTCTCAGCAGGCTGTTTATTGAAATGAAGCAGGAGGAATATGTTCAA 120

268 ACTGTGAAGTCATCAAAAGGTGTTCCGGATCAGCGGTAGCGCCCTATCTACCTTCAAT 327  
 121 ACTGTGAAGTCATCAAAAGGTGTTCCGGATCAGCGGTAGCGCCCTATCTACCTTCAAT 180

328 CCATCTCGGATGCTGCTGCTTCATAGGCCATATGTTAAAGGTGAGGATGAAGCA 387  
 181 CCATCTCGGATGCTGCTGCTTCATAGGCCATATGTTAAAGGTGAGGATGAAGCA 240

388 ACCATCATTCAGACTTCAACTAAGCGAAACAATGCAGCGTCAACAGATCAAAAGCAGCA 447  
 241 ACCATCATTCAGACTTCAACTAAGCGAAACAATGCAGCGTCAACAGATCAAAAGCAGCA 300

448 TATCTCAGGAAACAGAAAGCCCTGGATGAACACTGAAGAGCCCTTACAGTCA 507  
 301 TATCTCAGGAAACAGAAAGCCCTGGATGAACACTGAAGAGCCCTTACAGTCA 360

508 CTTGAGGAGGTGTTGTTTAGCTCTGCTAAAAAATCCAGCGCAATTTGATGCTGATGAAC 567  
 361 CTTGAGGAGGTGTTGTTTAGCTCTGCTAAAAAATCCAGCGCAATTTGATGCTGATGAAC 420

568 CGTGTGCTATGAAGGCGCTTGGAACTGATGAAGATCTCTAATGAGATTTTGGCATCA 627  
 421 CGTGTGCTATGAAGGCGCTTGGAACTGATGAAGATCTCTAATGAGATTTTGGCATCA 480

628 AGAACTAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGGAACCTGAAGAGAGAT 687  
 481 AGAACTAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGGAACCTGAAGAGAGAT 540

688 CTGGCCAAAGACATAA CTTGAGACACATCTGGAGATTTTCGGAAACGCTTTGCTTCTCT 747  
 541 CTGGCCAAAGACATAA CTTGAGACACATCTGGAGATTTTCGGAAACGCTTTGCTTCTCT 600

748 GCTAAGGTGACCGATCTGAGGACTTTGCTGATGAAGACTTGGCTGATTCAGATGCC 807  
 601 GCTAAGGTGACCGATCTGAGGACTTTGCTGATGAAGACTTGGCTGATTCAGATGCC 660

808 AGGGCCTTTGATGAAGCAGGAGAAAGGAGAAAGGGGACAGACGTAAACGTTTCAATACC 867  
 661 AGGGCCTTTGATGAAGCAGGAGAAAGGAGAAAGGGGACAGACGTAAACGTTTCAATACC 720

868 ATCTTACACAGAGACTTCCACACTTCGAGAGGTGTTTCAAGAAATACACCAAGTAC 927  
 721 ATCTTACACAGAGACTTCCACACTTCGAGAGGTGTTTCAAGAAATACACCAAGTAC 780

928 AGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGGTGACATTTGAGAAATGC 987  
 781 AGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGGTGACATTTGAGAAATGC 840

988 CTCACAGCTATCGTGAAGTGGCCACAGCAACACAGCTTTCTTTGAGAGAGCTTCAT 1047

841 CTCACAGCTATCGTGAAGTGGCCACCAAGCAACCAGCTTTCTTTGCAGAGAGCTTCAT 900  
 1048 CAAGCCATGAAGGTGTGGAACTCCGCCATPAAGGCATGATCAGGATATGTTTCCCGT 1107  
 901 CAAGCCATGAAGGTGTGGAACTCCGCCATPAAGGCATGATCAGGATATGTTTCCCGT 960  
 1108 TCTGAATTTGACATGAATGATATCAAAAGCATTTCTATCAGAAGATGATGATCTCCCTT 1167  
 961 TCTGAATTTGACATGAATGATATCAAAAGCATTTCTATCAGAAGATGATGATCTCCCTT 1020  
 1168 TGCCAAGCCATCTCGGATGAACCAAAAGGAGAGTATGAGAAAATCTTGGTGGCTCTTTGT 1227  
 1021 TGCCAAGCCATCTCGGATGAACCAAAAGGAGAGTATGAGAAAATCTTGGTGGCTCTTTGT 1080  
 1228 GGAGGAACCTAAACATTTCCCTTGGATGCTCAAGCTATGATCAGAAGACTTTTAATTATAT 1287  
 1081 GGAGGAACCTAAACATTTCCCTTGGATGCTCAAGCTATGATCAGAAGACTTTTAATTATAT 1140  
 1286 ATTTTCATCTTATTAAGCTTAAATAGSAAAGTTTCTTCAACAGGATTAACAGTGTAGCTACC 1347  
 1141 ATTTTCATCTTATTAAGCTTAAATAGSAAAGTTTCTTCAACAGGATTAACAGTGTAGCTACC 1200  
 1348 TACATGCTGAAAATATAGCCTTTAAATCATTTTATATATTAATTAATTAATTAATAGAGA 1407  
 1201 TACATGCTGAAAATATAGCCTTTAAATCATTTTATTAATTAATTAATTAATTAATAGAGA 1260  
 1408 TAAATCCATTTTTTAAAAATGTTTTTCCCAACCAATAAAAACCTTATACAAGTTTCTTAG 1467  
 1261 TAAATCCATTTTTTAAAAATGTTTTTCCCAACCAATAAAAACCTTATACAAGTTTCTTAG 1320  
 1468 TAAATCATGAGAGAGATGCTCTATGCTGCTGAAATTAATGCTGC 1516  
 1321 TAAATCATGAGAGAGATGCTCTATGCTGCTGAAATTAATGCTGC 1369

Search completed: March 10, 2004, 10:59:06  
 Job time : 373.965 secs



Db 346 TTTCCCGAGGACCTGCGAGTTTGTGAGATGACAGACGTGAAGGTCAACCATCATGTGAC 405  
QY 3053 CCGCTGAGAGTGAAGTGAACCGGCTACCGTGTGATGTATCCCCGTCAAACCTGCTGGC 3112  
Db 406 CCGCTGAGAGTGAACCGGCTACCGTGTGATGTATCCCCGTCAAACCTGCTGGC 465  
QY 3113 GAGCAGCGGAGAGGCTGCGCATGACAGGAACCTTTGCAGAGTCAACCGGCTGTCC 3172  
Db 466 GAGCAGCGGAGAGGCTGCGCATGACAGGAACCTTTGCAGAGTCAACCGGCTGTCC 525  
QY 3173 CCGGAGGTCACTTATTACTTCAAGTCTTTGCAAGTGAAGCATGAGAGGAGAGCAAGCT 3232  
Db 526 CCGGAGGTCACTTATTACTTCAAGTCTTTGCAAGTGAAGCATGAGAGGAGAGCAAGCT 585  
QY 3233 CTGACTGTCAACGAGACCAACCTGGAATGCTCCCACTAACCTCCAG-TTTGTCAATGA 3291  
Db 586 CTGACTGTCAACGAGACCAACCTGGAATGCTCCCACTAACCTCCAGTTTGTCAATGA 645  
QY 3292 AACTGATTCTACTGTCTGTGAGATGAGTCCACTCGGGCCCAAGATAACAGGATACCG 3351  
Db 646 AACTGATTCTACTGTCTGTGAGATGAGTCCACTCGGGCCCAAGATAACAGGATACCG 705  
QY 3352 ACTGACCTGGGCTTACCCGAAAGAGACAGCCAGGAGTAAATGTGGTCCCTGTGT 3411  
Db 706 ACTGACCTGGGCTTACCCGAAAGAGGACAGCCAGGAGTAAATGTGGTCCCTGTGT 765  
QY 3412 CTCGAGTACCACTGAGGAATCTGACGCTGACATCTGAGTACACCGTATCCCTGTGGC 3471  
Db 766 CTCGAGTACCACTGAGGAATCTGACGCTGACATCTGAGTACACCGTATCCCTGTGGC 825  
QY 3472 CATTAAGGGCAACCAAGAGAGCCCAAGCACTGAGTCTTTACCACTGACGCTGG 3531  
Db 826 CATTAAGGGCAACCAAGAGAGCCCAAGCACTGAGTCTTTACCACTGACGCTGG 885  
QY 3532 GAGCTATTTCACCTTACAAACCGAGGTGACTGAGACCACTTGTGATCAATGAC 3591  
Db 886 GAGCTATTTCACCTTACAAACCGAGGTGACTGAGACCACTTGTGATCAATGAC 945  
QY 3592 GCCT 3595  
Db 946 GCCT 949

Search completed: March 7, 2004, 21:39:02  
Job time : 16937 secs



```

/clones"CS0D1081V003"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 11.4%; Score 888; DB 13; Length 954;
Best Local Similarity 97.5%; Pred. No. 4e-244;
Matches 930; Conservative 0; Mismatches 21; Indels 3; Gaps 3;

2y 6334 CGTACCCACCTGGGTATGACACTGGAATGTTATTCAGCTTCCTGGCACTCTCGTCA 6393
2b 954 CGTACCCACCTGGGTATGACACTGGAATGTTATTCAGCTTCCTGGCACTCTCGTCA 895
2y 6394 GCAG-CCAGCTGTGGCAACAATGATCTTGG-AGGAACATGTTTATAGGCGGACCA 6451
2b 894 GCAGCCAGCTGTGGCGACAAATGATCTTGNAGNACATGTTTAGNTGGACCA 835
2y 6452 -CGCCCCACAAACCGCCACCCCATTAAGGCATAGGCCAAGACCATACCCGCGAATGAG 6510
2b 834 CGCCGCCACAAACCGCAGCCNCCATAGGCATAGGCCAAGACCATACCCGCGAATGAG 775
2y 6511 ACAGAGCTCTCTCTCAGACACCATCTCATGGGCCCATCCAGACACTCTCTAGTA 6570
2b 774 ACAGAGCTCTCTCTCAGACACCATCTCATGGGCCCATCCAGACACTCTCTAGTA 715
2y 6571 CATCATTTTCATGTCATCTCTGTCATGATGAGAACCCCTTACAGTTTCAGGTTCTCTG 6630
2b 714 CATCATTTTCATGTCATCTCTGTCATGATGAGAACCCCTTACAGTTTCAGGTTCTCTG 655
2y 6631 AACTTCTACAGTGCCTCTCTCAGAGCCCTCAGAGAGTGCACCTACACATCATCT 6690
2b 654 AACTTCTACAGTGCCTCTCTCAGAGCCCTCAGAGAGTGCACCTACACATCATCT 595
2y 6691 GGAGGCACTGAAGACACAGAGAGGATAGGTTTCGGAGAGAGTGTTCACCTGGGCA 6750
2b 594 GGAGGCACTGAAGACACAGAGGATAGGTTTCGGAGAGAGTGTTCACCTGGGCA 535
2y 6751 CTCTGTCAAGAGGCTTGAACCACTACAGTACTCTGTTGACCCCTTACACAGT 6810
2b 534 CTCTGTCAAGAGGCTTGAACCACTACAGTACTCTGTTGACCCCTTACACAGT 475
2y 6811 GTCCATTTATCCGCTTGGAGATGATGGGACGAATGCTCTGAATCAGGCTTTAACTGTT 6870
2b 474 TTCCATTTATCCGCTTGGAGATGATGGGACGAATGCTCTGAATCAGGCTTTAACTGTT 415
2y 6871 GTGCCAGTGTAGGCTTTGGAAGTGTCTATTTTCAGATGATTCATCTAGATGTTGCCA 6930
2b 414 GTGCCAGTGTAGGCTTTGGAAGTGTCTATTTTCAGATGATTCATCTAGATGTTGCCA 355
2y 6931 TGCAATGTTGTGAACCTACAGATTTGAGAGAGTGGGACCGTCAGGAGAAATGGCCA 6990
2b 354 TGCAATGTTGTGAACCTACAGATTTGAGAGAGTGGGACCGTCAGGAGAAATGGCCA 295
2y 6991 GATGATGAGTGCACATGCTCTTGGGAAACGGAAGGAGATTCAGTGTGACCCCTCATGA 7050
2b 294 GATGATGAGTGCACATGCTCTTGGGAAACGGAAGGAGATTCAGTGTGACCCCTCATGA 235
2y 7051 GGCAACGTTTACGATGATGGAGACATACCCAGTAGGAGACAGTGGCAGAGGAGATA 7110
2b 234 GGCAACGTTTATGATGATGGAGACATACCCAGTAGGAGACAGTGGCAGAGGAGATA 175
2y 7111 TCTCGGTGCCATTTGCTCCTCTGCATGCTTTTGGAGGCGAGCGGGCTGGCGCTGTGACAA 7170
2b 174 TCTCGGTGCCATTTGCTCCTCTGCATGCTTTTGGAGGCGAGCGGGCTGGCGCTGTGACAA 115
2y 7171 CTGGCGAGACTGGGGGTGAACCCAGTCCCGAGAGGCACTACTGGCCAGTCCCTACACCA 7230
2b 114 CTGGCGAGACTGGGGGTGAACCCAGTCCCGAGAGGCACTACTGGCCAGTCCCTACACCA 55

```

```

7231 GTATTTCTCAGAGATACCATCAGAGAACAAACACTAATGTTAATGTTGTTGTTG 7284
Db 54 GTATTTCTCAGAGATACCATCAGAGAACAAACACTAATGTTAATGTTGTTGTTG 1

RESULT 15
AL544172 949 bp mRNA linear EST 31-MAY-2003
LOCUS AL544172 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1020YB20 5-PRIME, mRNA sequence.
ACCESSION AL544172
VERSION AL544172.2 GI:31266017
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 949)
Li.W.B., Gruber,C., Jeess,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12876651.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1413.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1020DA10Q1&cluster=1413.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0D1020DA10Q1.
Location/Qualifiers
1..949
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1020YB20"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
Query Match 11.4%; Score 887.2; DB 9; Length 949;
Best Local Similarity 99.6%; Pred. No. 5.8e-244;
Matches 900; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

2693 ACAGGCTGTGATGCTTCTCTGACCCGACTGTGACCAAGTTGATGACCTCAATTGTT 2752
Db 46 ATAGGCTGTGATGCTTCTCTGACCCGACTGTGACCAAGTTGATGACCTCAATTGTT 105
2753 GTTCCCTGGACGAGACCCAGGCTCCCATACAGGTTACAGATAGTCTATTGCGCATCA 2812
Db 106 GTTCCCTGGACGAGACCCAGGCTCCCATACAGGTTACAGATAGTCTATTGCGCATCA 165
2813 GTAGAAGGTAGCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCTCAGT 2872
Db 166 GTAGAAGGTAGCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCTCAGT 225
2873 GACTTGCACCTGGTGTTCAGTATACATCTATCTATGCTGTGGAGAGAAATCAAGAA 2932
Db 226 GACTTGCACCTGGTGTTCAGTATACATCTATCTATGCTGTGGAGAGAAATCAAGAA 285
2933 AGTACACCTGTTGTTCATTCAACAGAAACCACTGGCACCCCAAGCTCAGATACAGTGGCC 2992
Db 286 AGTACACCTGTTGTTCATTCAACAGAAACCACTGGCACCCCAAGCTCAGATACAGTGGCC 345
2993 TCTCCAGGAGCTGCTGAGTTGTGGAGTACAGAGCTGAAGGTACCATCATGTGGACA 3052

```

DEFINITION AL551753 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CSOD1062YH09 5-PRIME, mRNA sequence.  
AL551753  
ACCESSION AL551753.2 GI:31273569  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1039)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 15, 2001 this sequence version replaced gi:12890006.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 1413.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSOD1062YH09&cluster=1413.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CSOD1062YH09P1.  
Location/Qualifiers  
1..1039  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSOD1062YH09"  
/tissue type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with NotI and cloned into the NotI and EcoRV  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 11.4%; Score 888.2; DB 9; Length 1039;  
Best Local Similarity 94.3%; Pred. No. 3.8e-244;  
Matches 931; Conservative 36; Mismatches 15; Indels 5; Gaps 5;

QY 6587 CTGTTGGACATGATGAAGAACCCCTACAGTTCAGGTTCTGMACTTCTACAGTGC 6616  
DB 55 CCGGGATGCACTGATGAAGAACCCCTACAGTTCAGGTTCTGMACTTCTACAGTGC 114  
QY 6647 ACTCTGACAGGCTCACCAGAGTGCCACCTACACATCATAGTGGAGGCACTGAAGAC 6706  
DB 115 ACTCTGACAGGCTCACCAGAGTGCCACCTACACATCATAGTGGAGGCACTGAAGAC 174  
QY 6707 CAGCAGAGCATTAAGTTTCGGGAAGAGGTTGTTACCGTGGGCACTCTGTCAAGAGGC 6766  
DB 175 CAGCAGAGCATTAAGTTTCGGGAAGAGGTTGTTACCGTGGGCACTCTGTCAAGAGGC 234  
QY 6767 TTGAACCACTACGATGACCTGCTGCTTACCCCTACACAGTCTCCATTATGCGCTT 6826  
DB 235 TTGAMCCAMCCTACGATGACCTGCTGCTTACCCCTACACAGTCTCCATTATGCGCTT 294  
QY 6827 GGAGATGAGTGGGAACGAATGTCTGAATCAGCGCTTAACTGTTGTGTCAGTGTAGGC 6886  
DB 295 GGAGATGAGTGGGAACGAATGTCTGAATCAGCGCTTAACTGTTGTGTCAGTGTAGGC 354  
QY 6887 TTGGAAGTGTCTATTTTCAGATGTGATTCATCTAGTGTGTCATGACATGGTGTGAAC 6946  
DB 355 TTGGAAGTGTCTATTTTCAGATGTGATTCATCTAGTGTGTCATGACATGGTGTGAAC 414  
QY 6947 TACAAGATTGGAGAGAGTGGGACCGTACAGGAGAAATGGCCAGATGATGAGCTGCACA 7006  
DB 415 TACAAGATTGGAGAGAGTGGGACCGTACAGGAGAAATGGCCAGATGATGAGCTGCACA 474  
QY 7007 TGTCTTGGGAACGGAAGAGGAAATTCAGTGTGACCTCATGAGGCAACGCTGTACGAT 7066

DB 475 TGTCTTGGGAACGGAAGAGGAAATTCAGTGTGACCTCATGAGGCAACGTTATGAT 534  
QY 7067 GATGGAAGACATACACAGTAGGAAACAGTGGCAGAGGAATATCTCGTGCCATTGTC 7126  
DB 535 GATGGAAGACATACACAGTAGGAAACAGTGGCAGAGGAATATCTCGTGCCATTGTC 594  
QY 7127 TCTGACACATGCTTTGGAGGCCAGCGGGCTGTGCAACTGCGGACAGACTGGG 7186  
DB 595 TCTGACACATGCTTTGGAGGCCAGCGGGCTGTGCAACTGCGGACAGACTGGG 654  
QY 7187 GGTGAACCCAGTCCGGAAGGCACTACTGGCCAGTCTCAACACAGTATTTCTCAGAGATAC 7246  
DB 655 GGTGAACCCAGTCCGGAAGGCACTACTGGCCAGTCTCAACACAGTATTTCTCAGAGATAC 713  
QY 7247 CATCAGAGAACAAACACTAATGTTAATGCCCCAATTTAGTGTCTCATGCTTTAGATGTA 7306  
DB 714 CATCAGAGAACAAACACTAATGTTAATGCCCCAATTTAGTGTCTCATGCTTTAGATGTA 773  
QY 7307 CAGGCTGACAGAGAGATTCGCCAGAGTAAATCATCTTTCCAAATCCAGAGGAACAAGCAT 7366  
DB 774 CAGGCTGACAGAGAGATTCGCCAGAGTAAATCATCTTTCCAAATCCAGAGGAACAAGCAT 833  
QY 7367 GTCTCTGCGCAAGATCCATCTAAACTGGAGAGTGTAGCAGACCCAGCTTAGATGTTCT 7426  
DB 834 GTCTCTGCGCAAGATCCATCTAAACTGGAGTGTAGCAGACCCAGCTTAGATGTTCT 892  
QY 7427 TCTTTCTTTCTTAAGCCCTTTCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCACAG 7486  
DB 893 TCTTTCTTTCTTAAGCCCTTTCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCACAG 951  
QY 7487 CTCTCTCAAGATCACCCTGGAGGTTCTCTGAGGTTTCTCATTAATGAGGCTGCACA 7546  
DB 952 CTCTCTCAAGATCACCCTGGAGGTTCTCTGAGGTTTCTCATTAATGAGGCTGCACA 1010  
QY 7547 TTGCTGTCTTCTGCTCGAAGTATTCAG 7573  
DB 1011 TTGCTGTCTTCTGCTCGAAGTATTCAG 1035

RESULT 14  
BX350259/c  
LOCUS BX350259  
DEFINITION BX350259 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CSOD1081Y003 3-PRIME, mRNA sequence.  
BX350259  
ACCESSION BX350259.1 GI:30375450  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 954)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 1413.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSOD1081Y003&cluster=1413.r.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CSOD1081Y003P1.  
Location/Qualifiers  
1..954  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

FEATURES  
source





Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 943)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 1413.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0BA10252G09\_CS02375\_1&cluster=1413.r.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0BA10252G09\_CS02375\_1.  
Location/Qualifiers

1. .943  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1065Y122"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

# ORIGIN

Query Match 11.6%; Score 901.2; DB 13; Length 943;  
Best Local Similarity 98.6%; Pred. No. 6.1e-248;  
Matches 928; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

648 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGCTCAGT 707  
649 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 880  
650 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 767  
651 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 820  
652 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 827  
653 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 760  
654 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 887  
655 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 700  
656 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 947  
657 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 641  
658 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 1007  
659 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 581  
660 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 1008  
661 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 521  
662 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 1127  
663 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 461  
664 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 1187  
665 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 401  
666 GCGCTGGGAGGGAATTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGT 1247

Db 400 ACTGTGTACAGACAGTGGTGGTCTACTCTGTGGGATGACAGTGGCTGAAGACACAG 341  
Qy 1248 GAAATTAAGCAAAATGCTTTGCAGTGGCTGGGCAACGAGATCAGTCCCAAGAGACAGCTG 1307  
Db 340 GAAATTAAGCAAAATGCTTTGCAGTGGCTGGGCAACGAGATCAGTCCCAAGAGACAGCTG 281  
Qy 1308 TAACCCAGACTTACGGTGGCACTCAATGGAGAGCCATGTCTTACCATTCACCTACA 1367  
Db 280 TAACCCAGACTTACGGTGGCACTCAATGGAGAGCCATGTCTTACCATTCACCTACA 221  
Qy 1368 ATGGCAGGAGCTTCTACTCTCTGCACACAGAGAGGGCGACAGGACGACATCTTTGGTGA 1427  
Db 220 ATGGCAGGAGCTTCTACTCTCTGCACACAGAGAGGGCGACAGGACGACATCTTTGGTGA 161  
Qy 1428 GCACAACTTCGAATATGAGCAGGACCAAGAAATACCTCTTCTGCAGACACCACTGTTT 1487  
Db 160 GCACAACTTCGAATATGAGCAGGACCAAGAAATACCTCTTCTGCAGACACCACTGTTT 101  
Qy 1488 TGGTTCAGACTCGAGGAGGAAATTCGAATGGTGCCTGTGGCCACTTCCCTTCTCTATACA 1547  
Db 100 TGGTTCAGACTCGAGGAGGAAATTCGAATGGTGCCTGTGGCCACTTCCCTTCTCTATACA 41  
Qy 1548 ACAACCAATTAACACTGATGATGACCTCTGAGGAGGAGAGA 1588  
Db 40 ACAACCAATTAACACTGATGATGACCTCTGAGGAGGAGAGA 1

RESULT 10  
AL541277  
LOCUS  
DEFINITION AL541277 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YN01  
S-PRIME, mRNA sequence.  
ACCESSION AL541277  
VERSION AL541277.2 GI:30545299  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 15, 2001 this sequence version replaced gi:12872191.  
Contact: Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 1413.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DE005CG01QPI&cluster=1413.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DE005CG01QPI.  
Location/Qualifiers

## FEATURES

source  
1. .1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE005YN01"  
/tissue\_type="PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo (dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

## ORIGIN

Query Match 11.5%; Score 900.2; DB 9; Length 1201;

```

QY 4775 GCTACATCAGCGGCTTAAACCTGGAGTTGATTATATACCATCACTGTGTGTCCTACT 4834
Db |||||
QY 808 GCTACATCAGCGGCTTAAACCTGGAGTTGATTATATACCATCACTGTGTGTCCTACT 867
Db |||||
QY 4835 GGCCTGGAGACAGCCCCGCGAAGCAGCAAGCAATTTCCATTAAATACCGAACAGAAATT 4894
Db |||||
QY 868 GGCCTGGAGACAGCCCCGCGAAGCAGCAAGCAATTTCCATTAAATACCGAACAGAAATT 927
QY 4895 GACAAACCATCCAGATGCAAGTGACCGATGTTACAGGACCAACAGCATTAGTGTCAAGTGG 4954
Db |||||
QY 928 GACAAACCATCCAGATGCAAGTGACCGATGTTACAGGACCAACAGCATTAGTGTCAAGTGG 986
QY 4955 CTGCTTCAAGTTCCTCTGTTACTGTTACAGGACCAACAGCATTCCAAAAAATGGACCA 5014
Db |||||
QY 987 CTGCTTCAAGTTCCTCTGTTACTGTTTACAGGACCAACAGCATTCCAAAAAATGGACAGG 1043
QY 5015 GGACCAACCAAAACTAAACCTGACAGTCCAGTCCAGA 5046
Db |||||
QY 1044 ACMAAAATAAATACTGCAGGTGTCGATCAACA 1075
Db |||||

RESULT 8
B0688596
LOCUS
DEFINITION B0688596 925 bp mRNA linear EST 15-JUL-2002
AGENCOURT_8341987 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6248842
5', mRNA sequence.
ACCESSION B0688596
VERSION B0688596.1 GI:21813912
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ARCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2389 row: f column: 11
High quality sequence stop: 659.
Location/Qualifiers
1..925
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6248842"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 11.6%; Score 905; DB 13; Length 925;
Best Local Similarity 99.4%; Pred. No. 4.8e-249;
Matches 919; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 5160 ACATTGATCGCCCTAAAGGACTGGGATTCACATGATGTGGATCGGATTCGATCCATCAAAATTG 5219
Db |||||

```

```

1 ACATTGATCGCCCTAAAGGACTGGCATTCACTGATGTGATGTGCGATTCCATCAAAATTG 60
5220 CTTGGGAAGACCCACAGGGGCAAGTTTCCAGGTACAGGTGACCTACTCGAGCCCTGAGG 5279
61 CTTGGGAAGACCCACAGGGGCAAGTTTCCAGGTACAGGTGACCTACTCGAGCCCTGAGG 120
5280 ATGGAATCCATGAGCTATTCCCTGCACTGTGATGTTGAAGAGACACTGAGAGTGCAG 5339
121 ATGGAATCCATGAGCTATTCCCTGCACTGTGATGTTGAAGAGACACTGAGAGTGCAG 180
5340 GCCTCAGACGGGTTCTGAGTACACAGTCAAGTGTGTTGCTTGCCTTGCACGATGATATG 5399
181 GCCTCAGACGGGTTCTGAGTACACAGTCAAGTGTGTTGCTTGCACGATGATATGAG 240
5400 GCCAGGCCCTGTGTTGAACCCAGTCCACAGCTATTCTCTGACCAACTGACCTGAAGTTC 5459
241 GCCAGGCCCTGTGTTGAACCCAGTCCACAGCTATTCTCTGACCAACTGACCTGAAGTTC 300
5460 CTCAGTCAACCCACAGCTGAGGCCAGTGGNACACCCATGTTCAAGTCTCACTG 5519
301 CTCAGTCAACCCACAGCTGAGGCCAGTGGNACACCCATGTTCAAGTCTCACTG 360
5520 GATATCGAGTGCGGGTGACCCCAAGGAGAACCCGACCAATGAAAGAAATCAACCTTG 5579
361 GATATCGAGTGCGGGTGACCCCAAGGAGAACCCGACCAATGAAAGAAATCAACCTTG 420
5580 CTCCTGACAGCTCATCCGTGTTGTTATCAGGACTTATGTGTCGCCACCAATATGAGTCA 5639
421 CTCCTGACAGCTCATCCGTGTTGTTATCAGGACTTATGTGTCGCCACCAATATGAGTCA 480
5640 GTCTCTATGCTCTTAAGGACACTTTGACAAAGCAGACCACTCAGGAGTGTGTCAACCTC 5699
481 GTCTCTATGCTCTTAAGGACACTTTGACAAAGCAGACCACTCAGGAGTGTGTCAACCTC 540
5700 TGAGAAATGTCAGCCCAACAAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAATCA 5759
541 TGAGAAATGTCAGCCCAACAAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAATCA 600
5760 CCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAAAGTTGATGCGTTCCAG 5819
601 CCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAAAGTTGATGCGTTCCAG 660
5820 CCAATGCCAGACTTCCAATCCAGAGAACCAATCAAGCCAGATGTGAGAACTCACCATCA 5879
661 CCAATGCCAGACTTCCAATCCAGAGAACCAATCAAGCCAGATGTGAGAACTCACCATCA 720
5880 CTCGCTTACAAACAGGACTGACTACAGATCTACTGTACACCTTGAATGATCAATGCTC 5939
721 CAGGTTTACAAACAGGACTGACTACAGATCTACTGTACACCTTGAATGATCAATGCTC 780
5940 GGAGCTCCCCCTGTGTGTCATCGAGCCCTCCACTGCCATTGATGACCATCCAACTCGTT 5999
781 GGAGCTCCCCCTGTGTGTCATCGAGCCCTCCACTGCCATTGATGACCATCCAACTCGTT 840
6000 TCTGCGCACCAACACCAATTCCTGCTGATCATGGAGCGGCCACCGTCCAGG-ATT 6058
841 TCTGCGCACCAACACCAATTCCTGCTGATCATGGAGCGGCCACCGTCCAGG-ATT 900
6059 ACCGGCTACATCATCAAGTATGAGA 6083
901 ACCGGCTACATCATCAAGTATGAGA 925

RESULT 9
BX391752/c
LOCUS
DEFINITION BX391752 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1065Y122 3-PRIME, mRNA sequence.
ACCESSION BX391752
VERSION BX391752.1 GI:30607701
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```



```

488 GGAAGAACATCATCAGATATAATCAACAGTGGGAGCGGACCTACCTAGGCAATGCGTGTGTT 547
494 TGTACTTGTATGGAGAACCGAGGTTTAACTCGAAAGTAAACCTGAAGCTGAAGAG 553
548 TGTACTTGTATGGAGAACCGAGGTTTAACTCGAGAGTAAACCTGAAGCTGAAGAG 607
554 ACTTGCTTTGACAGTACACTGGGACACTTACCGAGTGGGTGACACTTATGAGGTCCT 613
608 ACTTGCTTTGACAGTACACTGGGACACTTACCGAGTGGGTGACACTTATGAGGTCCT 667
614 AAAGACTCCATGATCTGGGACTGTACTCTGCATCGGGGCTGGCGGAGGAGATAAGCTGT 673
668 AAAGACTCCATGATCTGGGACTGTACTCTGCATCGGGGCTGGCGGAGGAGATAAGCTGT 727
674 ACCATGCGAAACCGCTGCCATGAGGGGTGAGTCTCAAGATTTGCTGACACCTGGAGG 733
728 ACCATGCGAAACCGCTGCCATGAGGGGTGAGTCTCAAGATTTGCTGACACCTGGAGG 787
734 AGACCATGAGACTGGTGTGTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAGGA 793
788 AGACCATGAGACTGGTGTGTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAGGA 847
794 GAATGAGCTGCAAGCCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACTTCCTAT 853
848 GAATGAGCTGCAAGCCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACTTCCTAT 907
854 GTGGTGGGAAACCGCTGCCATGAGGGGTGAGTCTCAAGATTTGCTGACACCTGGAGG 913
908 GTGGTGGGAAACCGCTGCCATGAGGGGTGAGTCTCAAGATTTGCTGACACCTGGAGG 967
914 CTGGGAGAGCGGAGCGCATCTGTCATCTTCAAGTGTGATGATGATGATGATGATGATG 973
968 CTGGGAGAGCGGAGCGCATCTGTCATCTTCAAGTGTGATGATGATGATGATGATGATG 1026
974 ACAAGGACATCTTATAGATTTGGAGACCTCGGAGCAAGAGGATAATCGAGAAACCTG 1033
1027 ACAGG-----ACATCTATAGATTTGGGMACTGGAGCAAGAGATGATGATGATGATG 1082
1034 CTCAGTGCATCTGCACAGCAACCGCGGAGGAGAGTGAAGTGTGAGGAGCAGACCTCT 1093
1083 GTGCAYVYKMGAGCAGCGGCGGGGTGGAGTGTGRRGGMCTCTTGMGACWYACRGS 1142
1094 GTCGAGACC 1102
1143 GRTYTGSCC 1151

RESULT 7
BX439746 1201 bp mRNA linear EST 15-MAY-2003
LOCUS BX439746 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE014YA12
DEFINITION S-PRIME, mRNA sequence.
ACCESSION BX439746
VERSION BX439746.1 GI:30773802
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1201),
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1413.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE014BA06QPI&cluster=1413.r. Contact :
Peng Liang Email : fliang@lifetech.com URL :

```

```

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DE014BA06QPI.
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE014YA12"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 11.9%; Score 924; DB 13; Length 1201;
Best Local Similarity 94.4%; Pred. No. 1.9e-254;
Matches 993; Conservative 7; Mismatches 47; Indels 5; Gaps 4;

3996 ACAATGTCAGTGTTCACATGTCACAGGATGACAGGAAAGTGTCCCTATCTCTGATACCA 4055
28 AAAAAACAGCGTGGTACCGGTCGGGAATTCGGGATAGTGTCCCTATCTCTGATACCA 87
4056 TCATCCAGTGTTCCTCCCTCCCACTGACCTGCGATTCACCAACATTTGGTCCAGACCA 4115
88 TCATCCAGTGTTCCTCCCTCCCACTGACCTGCGATTCACCAACATTTGGTCCAGACCA 147
4116 TGCGTGTACCTGGGTCACCCCATCCATTCATTTAAACCAACTTCCTGGTGGTACT 4175
148 TGCGTGTACCTGGGTCACCCCATCCATTCATTTAAACCAACTTCCTGGTGGTACT 207
4176 CACTGTGAAATAGAGAAAGATTTGACAGATTTGCAATTTCTCTTCAGACAATGCAG 4235
208 CACTGTGAAATAGAGAAAGATTTGACAGATTTGCAATTTCTCTTCAGACAATGCAG 267
4236 TGGTCTTAAACAAATCTCTGCTGGTACAGAAATATAGTAGTGTCTCCAGTGTCTACG 4295
268 TGGTCTTAAACAAATCTCTGCTGGTACAGAAATATAGTAGTGTCTCCAGTGTCTACG 327
4296 AACCAATGAGAGACACCTCTTAGAGAAAGAGAGAAAAGTGTTCCTGATTCCTCAACTG 4355
328 AACCAATGAGAGACACCTCTTAGAGAAAGAGAGAAAAGTGTTCCTGATTCCTCAACTG 387
4356 GCATTGACTTTCTGATATTACTGCCCACTCTTTTACTGTGCACTGATTCCTCGAG 4415
388 GCATTGACTTTCTGATATTACTGCCCACTCTTTTACTGTGCACTGATTCCTCGAG 447
4416 CCACCATCACTGGCTACAGGATCCGCCATCATCCGAGCACTTCAGTGGAGACCTCGAG 4475
448 CCACCATCACTGGCTACAGGATCCGCCATCATCCGAGCACTTCAGTGGAGACCTCGAG 507
4476 AAGATCGGGTCCCACTCTCGGAATTCCTCACTCCCACTTCAGTGGAGACCTCGAG 4535
508 AAGATCGGGTCCCACTCTCGGAATTCCTCACTCCCACTTCAGTGGAGACCTCGAG 567
4536 AGTATGTGGTTCAGCATCGTTCCTCTTAATGSCAGAGGAAAGTTCCTTATTTGGCC 4595
568 AGTATGTGGTTCAGCATCGTTCCTCTTAATGSCAGAGGAAAGTTCCTTATTTGGCC 627
4596 AACATCAACAGTTCATGATTCGAGGAGACCTGGAGTGTGCTGCGACCCCACTCA 4655
628 AACATCAACAGTTCATGATTCGAGGAGACCTGGAGTGTGCTGCGACCCCACTCA 687
4656 GCCTACTGATCAGCTGGGATGCTCTGCTGTACAGTGAATATTACAGGATCACTTACG 4715
688 GCCTACTGATCAGCTGGGATGCTCTGCTGTACAGTGAATATTACAGGATCACTTACG 747
4716 GAGAAACAGGAGAAATAGCCCTGTCCAGGAGTTCAGTGTGCTTGGAGCACTTACA 4774
748 GAGAAACAGGAGAAATAGCCCTGTCCAGGAGTTCAGTGTGCTTGGAGCACTTACA 807

```

2y	11	CGCGCTGTGTCACACAGGGGAGAGAGGNA	CCCACAGGCGCGAGCGGAGAGGGGACCT	70
Db	66	CGCGCTGTGTCACACAGGGGAGAGAGGNA	CCCCAGGCGCGAGCGGAGAGGGGACCT	125
2y	71	GCAGCCACAACTTCTCTGGTCTCTGCATCCCTTCTGTCCCTCCACCCGCTCCCGCTTCCCGC	130	
Db	126	GCAGCCACAACTTCTCTGGTCTCTGCATCCCTTCTGTCCCTCCCHCCCGTCCCGCTTCCCGC	185	
2y	131	ACCGTCTGGCCCCACGTTCTTGGAGGCGACCAACCCCGGGAGGCAATTAGAGGGGATTTT	190	
Db	186	ACCGTCTGGCCCCACGTTCTTGGAGGCGACCAACCCCGGGAGGCAATTAGAGGGGATTTT	245	
2y	191	TCCCGCAGGTTGCGAAGGGAAGCAAACTTGGTGGCAACTTGGCTCCCGTTCGCGGCGGTCT	250	
Db	246	TCCCGCA-GTTTGGAGGGAAGCAAACTTGGTGGCAACTTGGCTCCCGTTCGCGGCGGTCT	304	
2y	251	CTCCCCACCGTCTCAACATGCTTAGGGGTCCGGGGCCCGGGCTGTCTGTCTGGCGGTG	310	
Db	305	CTCCCCACCGTCTCAACATGCTTAGGGGTCCGGGGCCCGGGCTGTCTGTCTGGCGGTG	364	
2y	311	CTGTGCTGGGACAGCGGTGCCCTCCACGGGAGCCTCGAAGAGCAAGAGCGCAGGCTCAG	370	
Db	365	CAGTGCTGGGACAGCGGTGCCCTCCACGGGAGCCTCGAAGAGCAAGAGCGCAGGCTCAG	424	
2y	371	CAAAATGGTTGAGCCCCAGTCCCCGGTGGCTGTGAGTCAAGAGCAAGCCCGTGTGTTATGAC	430	
Db	425	CAAAATGGTTGAGCCCCAGTCCCCGGTGGCTGTGAGTCAAGAGCAAGCCCGTGTGTTATGAC	484	
2y	431	AATGGAAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGCAATGCGTTG	490	
Db	485	AATGGAAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGCAATGCGTTG	544	
2y	491	GTATTGACTTGTATTGGAGGAAGCCGAGGTTTTAACTGCGAAAGTAAACCTGAAGCTGAA	550	
Db	545	GTATTGACTTGTATTGGAGGAAGCCGAGGTTTTAACTGCGAGAGTTAACTGAAGCTGAA	604	
2y	551	GAGACTTGCTTTGACAAGTACACTTGGGAAACACTTACCGAGTGGGTGACACTTATGAGCGT	610	
Db	605	GAGACTTGCTTTGACAAGTACACTTGGGAAACACTTACCGAGTGGGTGACACTTATGAGCGT	664	
2y	611	CCTAAAGACTCCATGATCTGGGACTGTACCTGTGCATCGGGCTGGGCGAGGGAATAAGC	670	
Db	665	CCTAAAGACTCCATGATCTGGGACTGTACCTGTGCATCGGGCTGGGCGAGGGAATAAGC	724	
2y	671	TGTACCATCCGAACCGCTCCATGAAGGGGGTCAGTCCCTACAAGATTGTTGTGACACCTGG	730	
Db	725	TGTACCATCCGAACCGCTCCATGAAGGGGGTCAGTCCCTACAAGATTGTTGTGACACCTGG	784	
2y	731	AGGAGACCACATGAGACTGTGGTTACATGTTAGAGTGTGTGTCTGTGTAATGGAAA	790	
Db	785	AGGAGACCACATGAGACTGTGGTTACATGTTAGAGTGTGTGTCTGTGTAATGGAAA	844	
2y	791	GGAGAAATGGACCTCGAAGCCATAGCTGAGAAGTGTGTTTGCATCATGCTGCTGGGACTTCC	850	
Db	845	GGAGAAATGGACCTCGAAGCCATAGCTGAGAAGTGTGTTTGCATCATGCTGCTGGGACTTCC	904	
2y	851	TATGTGCTCGAGAAAACGTTGGGAAAGCCCTACCAAGCGTGGATGATGTTGATGTTGACT	910	
Db	905	TATGTGCTCGAGAAAACGTTGGGAAAGCCCTACCAAGCGTGGATGATGTTGATGTTGACT	964	
2y	911	TGCGTGGAGNAGCGCGGACGATCACTTGCACTTCTAGAAATAGATCAACAGATCAG	970	
Db	965	TGCGTGGAGNAGCGCGGACGATCACTTGCACTTCTAGAAATAGATCAACAGATCAG	1024	
2y	971	GACACAAGGACATCCTATAGAAATGGAG	998	
Db	1025	GACACAAGGACATCCTATAGAAATGGAG	1052	

[illegible]

5-PRIME, mRNA sequence.  
BX417945.1 GI:30642158  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1201)  
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 1413.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DE014BG12QP1&cluster=1413.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0DE014BG12QP1.  
Location/Qualifiers  
1. .1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE014YM24"  
/tissue\_type="PLACENTA"  
/clone\_lib="Homo sapiens PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-cligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

ORIGIN

Query Match 12.2%; Score 947.8; DB 13; Length 1201;  
Best Local Similarity 91.7%; Pred. No. 2.4e-261;  
Matches 999; Conservative 21; Mismatches 64; Indels 5; Gaps 2;

QY	14	CTGTGTCGACAGGGGGAGGAGGGAA	CCC	CAGGCGCGAGCGGAGAGGGGACCTG	GCA	73
Db	68	CTGTGTCGACAGGGGGAGGAGGGAA	CCC	CAGGCGCGAGCGGAGAGGGGACCTG	GCA	127
QY	74	GCACAACTTCTGTGTCTCTGTCA	TCCCTTCGTCCTCCACCCGCTCC	CCCTTCCTCCCAACC	133	
Db	128	GCACAACTTCTGTGTCTCTGTCA	TCCCTTCGTCCTCCCTCCGCTTC	CCCTTCCTCCCAACC	187	
QY	134	CTCTGGCCCCACCTCTTTGGAGGCG	CACAAACCCCGGGAGGCAATTAGA	AGGGATTTTTTC	193	
Db	188	CTCTGGCCCCACCTCTTTGGAGGCG	CACAAACCCCGGGAGGCAATTAGA	AGGGATTTTTTC	247	
QY	194	CGCAGGTTCGGAAGGGAAGCAAA	CTTTGGTGGCAACTTGCTCCCGGTG	CGGGCGTCTCTC	253	
Db	248	CGCAGGTTCGGAAGGGAAGCAAA	CTTTGGTGGCAACTTGCTCCCGGTG	CGGGCGTCTCTC	307	
QY	254	CCCAACCGTCTCAACATGCTTAG	GGGTCCCGGGGCCCGGGCTGTGCTG	CTGGCCGTCCTG	313	
Db	308	CCCAACCGTCTCAACATGCTTAG	GGGTCCCGGGGCCCGGGCTGTGCTG	CTGGCCGTCCTG	367	
QY	314	TGCCTGGGGACAGGGTGCCTCC	ACGGGAGCCTCGAAGACG	AAGACGAGGCGTCAGCAA	373	
Db	368	TGCCTGGGGACAGGGTGCCTCC	ACGGGAGCCTCGAAGACG	AAGACGAGGCGTCAGCAA	427	
QY	374	ATGTTTCAGCCCCAGTCCCGGTGG	TGTGCAGTCAAAAGCAAGCCCGGT	TGTTATGACAAT	433	
Db	428	ATGTTTCAGCCCCAGTCCCGGTGG	TGTGCAGTCAAAAGCAAGCCCGGT	TGTTATGACAAT	487	
QY	434	GGAAAAACATTCAGATTAATCA	ACAGTGGGAGCGGACCTACCT	PAGSCAATGCGTTGGTT	493	

cgi-bin/cluster.cgi?seq=CS0DE008BH010P1&cluster=1413.r. Contact :  
 Feng Liang Email: fliang@lifetech.com URL: Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DE008BH010P1.  
 Location/Qualifiers  
 1. .1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE008Y002"  
 /tissue\_type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."  
 ORIGIN  
 Query Match 13.4%; Score 1043.2; DB 9; Length 1201;  
 Best Local Similarity 95.8%; Pred. No. 6.3e-289;  
 Matches 1086; Conservative 24; Mismatches 19; Indels 5; Gaps 4;  
 6357 CTGGAAATGGTATTCAGCTTCCTGGCACTCTGTCAGCAACCCAGTGTGGCAACAAA 6416  
 57 CTGGAAATGGTATTCAGCTTCCTGGCACTCTGTCAGCAACCCAGTGTGGCAACAAA 116  
 6417 TGATCTTTGAGGAACATGTTTTAGCGGACCAACCGCCCAACCGCCACCCCATAA 6476  
 117 TGATCTTTGAGGAACATGTTTTAGCGGACCAACCGCCCAACCGCCACCCCATAA 176  
 6477 GGCATAGGCGAAGACATACCCCGCAAGTAGGACAGAAAGCTCTCTCAGCAACCA 6536  
 177 GGCATAGGCGAAGACATACCCCGCAAGTAGGACAGAAAGCTCTCTCAGCAACCA 236  
 6537 TCTCATGGGCCCCATTCAGGACACTTCTGAGTACATCATCTTCATCTCTGTGGCA 6596  
 237 TCTCATGGGCCCCATTCAGGACACTTCTGAGTACATCATCTTCATCTCTGTGGCA 296  
 6597 CTGATGAAGAACCTTTACAGTTGAGGTTCTGGAACTTCTACCACTGACATCTGACAG 6656  
 297 CTGATGAAGAACCTTTACAGTTGAGGTTCTGGAACTTCTACCACTGACATCTGACAG 356  
 6657 GCCTCACCAGAGTGCACCTACACATCATAGTGGAGCACTGAAAGACCAAGAGAGGC 6716  
 357 GCCTCACCAGAGTGCACCTACACATCATAGTGGAGCACTGAAAGACCAAGAGAGGC 416  
 6717 ATAAGTTGCGGAAGAGGTTGTTACCGTGCGCAACTCTCTCAACGAAGCTTTGAACCAAC 6776  
 417 ATAAGTTGCGGAAGAGGTTGTTACCGTGCGCAACTCTCTCAACGAAGCTTTGAACCAAC 476  
 6777 CTACGGATGACTCGTCTTTGACCCCTACACAGTGTCCCATTTATGCGTTGGAGATGAGT 6836  
 477 CTACGGATGACTCGTCTTTGACCCCTACACAGTGTCCCATTTATGCGTTGGAGATGAGT 536  
 6837 GGGAAACGAATGCTGAATCAGGCTTTAAACTGTTGTCAGTGTCTTGGAGTTGAGTGTG 6896  
 537 GGGAAACGAATGCTGAATCAGGCTTTAAACTGTTGTCAGTGTCTTGGAGTTGAGTGTG 596  
 6897 GTCAATTCAGATGATTCATCTAGATGGTCCATGACATGTTGTGTAACCTCAAGATTG 6956  
 597 GTCAATTCAGATGATTCATCTAGATGGTCCATGACATGTTGTGTAACCTCAAGATTG 656  
 6957 GAGAGAAGTGGGACCGCTCAGGGAAGAAATGGCCAGATGATGAGCTGCAATGCTTTGGGA 7016  
 657 GAGAGAAGTGGGACCGCTCAGGGAAGAAATGGCCAGATGATGAGCTGCAATGCTTTGGGA 715  
 7017 ACGGAAGAGGATTCAGTGTGACCTCATCAGGCAAGCTGTTACGATGATGGGAAGA 7076  
 716 ACGGAAGAGGATTCAGTGTGACCTCATCAGGCAAGCTGTTACGATGATGGGAAGA 775  
 7077 CATACCAAGTGGAGAGACAGTGGCAGAAAGAAATATCTCGTGCCATTTCTCTGACAT 7136

776 CATACCAAGTAGGAGAACAGTGGCAGAGGAATATCTCGTGGCCATTTGCTCTGACAT 835  
 7137 GCCTTTGAGGCGCAGCGGCTGCGCTGTGACAACCTGCCGACCTGCGGCTGAACCCA 7196  
 836 GCCTTTGAGGCGCAGCGGCTGCGCTGTGACAACCTGCCGACCTGCGGCTGAACCCA 895  
 7197 GTCCGAAAGCACTACTGCGGCACTCTCAACACAGTATTTCTCAGAGATACCATCAGAGAA 7256  
 896 GTVCCGAA-GCACTACTGCGGCACTCTCAACACAGTATTTCTCAGAGATACCATCAGAGAA 954  
 7257 CAACACTAATGTTAATGCCCCAATGAGTGTTCATGCTTTAGATGTACAGGCTGACA 7316  
 955 CAACACTAATGTTAATGCCCCAATGAGTGTTCATGCTTTAGATGTACAGGCTGACA 1014  
 7317 GAGAAGATTCCTCGAGAGTAAATCATCTTTCCAAATCCAGAGGAACAAGCATGTCTCTGC 7376  
 1015 GAGAAGATTCCTCGAGAGTAAATCATCTTTCCAAATCCAGAGGAACAAGCATGTCTCTGC 1074  
 7377 CAGATCCATCTAATGAGTGTAGTGTAGCAGACCCAGCTTAGATCTTCTTCTTCTTC 7436  
 1075 CAGATCCATCTAATGAGTGTAGTGTAGCAGACCCAGCTTAGATCTTCTTCTTCTTCTT 1133  
 7437 TTAAGCCCTTTGCTCTGAGAGAGTCTTCCAGCTTCAAGCTCAATCAGAGCTTC 7490  
 1134 TWAR--CCCTTTCVCKGGGGKYYVCMRSYTTAGSCCAHYACARGBTTC 1185  
 RESULT 5  
 BX386270 1052 bp mRNA linear EST 08-MAY-2003  
 LOCUS BX386270 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS0DI058YA10 5-PRIME, mRNA sequence.  
 ACCESSION BX386270  
 VERSION BX386270.1 GI:30447433  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1052)  
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 JOURNAL Contact: Genoscope  
 COMMENT Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 1413.r For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS1AI015ZC06QP1&cluster=1413.r. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS1AI015ZC06QP1.  
 FEATURES  
 Location/Qualifiers  
 1. .1052  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DI058YA10"  
 /tissue\_type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoRV  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 ORIGIN  
 Query Match 12.4%; Score 970; DB 13; Length 1052;  
 Best Local Similarity 99.4%; Pred. No. 8.4e-268;  
 Matches 982; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

2Y	6017	AAATCCTTGCTGTGTATCATGGCAGCGCCACGTCGCGAGGATTACCGGCTACATCATCAAG	6076
Db	311	AAATTCCTTGCTGTGTATCATGGCAGCGCCACGTCGCGAGGATTACCGGCTACATCATCAAG	370
2Y	6077	TATGAGAGCCTGGGTCTCCTCCACGAGAACTGGTCCCTCGGCGCCGCTGTGTGCACA	6136
Db	371	TATGAGAGCCTGGGTCTCCTCCACGAGAACTGGTCCCTCGGCGCCGCTGTGTGCACA	430
2Y	6137	GAGGCTACTATTACTGGCTTGGAAACCGGGAAACGGAATATACAAATTTATGTCAATGTCGCTTG	6196
Db	431	GAGGCTACTATTACTGGCTTGGAAACCGGGAAACGGAATATACAAATTTATGTCAATGTCGCTTG	490
2Y	6197	AGAAATAATCAGAAAGACGAGCCCTGTATTTGGAAAGGAAAGAAAGACAGACGAGCTTCCCCAA	6256
Db	491	AGAAATAATCAGAAAGACGAGCCCTGTATTTGGAAAGGAAAGAAAGAC	534
2Y	6257	CTGGTAACCCCTTCCACACCCCAATCTTCAATGGACCGAGAGATCTTGGATGTCCTTCCACA	6316
Db	535	-----	534
QY	6317	GTTCAAAAGACCCCTTTCGTCAACCACCTGGGTATGACACTGGAAATGGTATTCAGCTT	6376
Db	535	-----	534
QY	6377	CCTGGCACCTTCTGGTCAGCAACCCAGTGTGGGCACAACAATGATCTTTGAGGAACATGTT	6436
Db	535	-----	534
2Y	6437	TTTAGGGGACCAACCGCCCAACAAACCGCCACCCCAATAGGCATAGGCCAAGACCATAC	6496
Db	535	-----	534
QY	6497	CCGCCGAATGTAGACAAAGAGCTCTCTCAGACAAACCATTCTCATGGGCGCCCATTCACG	6556
Db	535	-----AGGACAAAGAGCTCTCTCTCAGACAAACCATTCTCATGGGCGCCCATTCACG	583
QY	6557	GACACTTCTCAGTACATCAATTTCAATGTCATCTCTTGGCACTGATGAAGAAACCCCTTACAG	6616
Db	584	GACACTTCTCAGTACATCAATTTCAATGTCATCTCTTGGCACTGATGAAGAAACCCCTTACAG	643
QY	6617	TTCAGGGTTCCTGGAACTTCTACCAAGTGCCACTCTGACAGGCTCACCAGAGGTGCCACC	6676
Db	644	TTCAGGGTTCCTGGAACTTCTACCAAGTGCCACTCTGACAGGCTCACCAGAGGTGCCACC	703
QY	6677	TACAACATCATAGTGGAGGCACTGAAAGACACAGCAGAGGCATTAAGGTTCCGGAAGAGGTT	6736
Db	704	TACAACATCATAGTGGAGGCACTGAAAGACACAGCAGAGGCATTAAGGTTCCGGAAGAGGTT	763
QY	6737	GTTACCGTGGCAACTCTGTCAACGAGAGGCTTGAACCAACCTACGGATGACTCGTGCTTT	6796
Db	764	GTTACCGTGGCAACTCTGTCAACGAGAGGCTTGAACCAACCTACGGATGACTCGTGCTTT	823
QY	6797	GACCCCTACACAGTGTCCCATTTATGCGGTTGGAGATGAGTGGGAACGAAATGTCGATCA	6856
Db	824	GACCCCTACACAGTGTCCCATTTATGCGGTTGGAGATGAGTGGGAACGAAATGTCGATCA	883
QY	6857	GGCTTTAAACGTGTTGGCCAGTGCTTAGGCTTTGGAAAGTGGTCATTTTCAGATGTGATTC	6916
Db	884	GGCTTTAAACGTGTTGGCCAGTGCTTAGGCTTTGGAAAGTGGTCATTTTCAGATGTGATTC	943
QY	6917	TCTAGATGGTCCCATGACAAATGGTGTGAACCTACAAGATTGAGAGAAAGTGGACCGTCAG	6976
Db	944	TCTAGATGGTCCCATGACAAATGGTGTGAACCTACAAGATTGAGAGAAAGTGGACCGTCAG	1003
QY	6977	GGAGAAATGSCCAAGATGATGAGCTGCAATGTCTTGGAAACGGAAAGAGGAATTCAG	7036
Db	1004	GGAGAAATGSCCAAGATGATGAGCTGCAATGTCTTGGAAACGGAAAGAGGAATTCAG	1063
QY	7037	TGTGACCTCATGAGGCAACCTGTTACGATCATGGAGACATACCACGTAGGAGAACAG	7096
Db	1064	TGTGACCTCATGAGGCAACCTGTTACGATCATGGAGACATACCACGTAGGAGAACAG	1123
QY	7097	TGGCAGAGGAATATCTCGGTGCCATTTGCTCCTGCACATGCTTTTGGAGGCCACGGGGC	7156

1124	Db	TTGGCAGAGGAATATCTCGTGCCATTCTCTGCACATGCTTTGAGGCCAGCGGGC	1183
7157	Qy	TGGCGCTGTGACAACTGCCGACAGACTGGGGGTGAACCCAGTCCCGAAGGCACTACTGGC	7216
1184	Db	TGGCGCTGTGACAACTGCCGACAGACTGGGGGTGAACCCAGTCCCGAAGGCACTACTGGC	1243
7217	Qy	CAGTCTCTCAACCCAGTATTCTCAGAGATACCATTACAGAGAACAACACTAATGTTAATTGC	7276
1244	Db	CAGTCTCTCAACCCAGTATTCTCAGAGATACCATTACAGAGAACAACACTAATGTTAATTGC	1303
7277	Qy	CCAATTGAGTGTCTCATGCTTTAGATGTACAGGCTGACAGAGAAGATTCCCGAGAGTAA	7336
1304	Db	CCAATTGAGTGTCTCATGCTTTAGATGTACAGGCTGACAGAGAAGATTCCCGAGAGTAA	1363
7337	Qy	ATCATCTTTCCCAATCCAGAGGAACAAGCATGCTCTCTGCCAAGATCCATCTAAACTGGA	7396
1364	Db	ATCATCTTTCCCAATCCAGAGGAACAAGCATGCTCTCTGCCAAGATCCATCTAAACTGGA	1423
7397	Qy	GTGANTTTAGCAGACCCAGCTTAGAGTTCTTCTTTTAAAGCCCTTGTCTCTGGAG	7456
1424	Db	GTGANTTTAGCAGACCCAGCTTAGAGTTCTTCTTTTAAAGCCCTTGTCTCTGGAG	1483
7457	Qy	GAGATTCTCCAGCTTCAGCTCAACTCACAGCTCTCCAAAGCATCACCTGGGAGTTTCCT	7516
1484	Db	GAGATTCTCCAGCTTCAGCTCAACTCACAGCTCTCCAAAGCATCACCTGGGAGTTTCCT	1543
7517	Qy	GAGGTTTTCTCATAAATGAGGGCTGCACATTGCCCTGTCTGCTTCGAAGTATTCAATAC	7576
1544	Db	GAGGTTTTCTCATAAATGAGGGCTGCACATTGCCCTGTCTGCTTCGAAGTATTCAATAC	1603
7577	Qy	CGCTCAGTATTTTAAATGAAGTGATTCTAAGATTTCGTTTGGGATCAATAGGAAAGCATA	7636
1604	Db	CGCTCAGTATTTTAAATGAAGTGATTCTAAGATTTCGTTTGGGATCAATAGGAAAGCATA	1663
7637	Qy	TGCAGGCCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAAATTTTAAAGTAGGAAAG	7696
1664	Db	TGCAGGCCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAAATTTTAAAGTAGGAAAG	1723
7697	Qy	TCACCCCAAACACTTCTGCTTTCACTTTAAGTGCTGGCCGCCCAATACTGTAGGAAACAAGCA	7756
1724	Db	TCACCCCAAACACTTCTGCTTTCACTTTAAGTGCTGGCCGCCCAATACTGTAGGAAACAAGCA	1783
7757	Qy	TGATCTTGTTACTGTGATATTTTAAATATCCACAGTACT	7795
1784	Db	TGATCTTGTTACTGTGATATTTTAAATATCCACAGTACT	1822

RESULT 4	ACCESSION	REFERENCE
ALS42212	VERSION	AUTHORS
LOCUS	KEYWORDS	TITLE
DEFINITION	SOURCE	JOURNAL
	ORGANISM	COMMENT

AL542212 1201 bp mRNA linear EST 12-MAY-2003  
AL542212 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE008Y002  
S-PRIME, mRNA sequence.  
AL542212  
AL542212.2 GI:30547139  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1201)  
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 15, 2001 this sequence version replaced gi:13874034.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 1413.r For  
more information about this cluster, see  
<http://www.genoscope.cns.fr/>

```
6508 -----AGGACAGAAGCT 6520
b TACCACCTCTATCTCTCATGTTCCGGGCTCAATCCAAATGCTCTACAGGACAGGAAGCT 6819
y CTCTCTCAGACACCACTCTCATGGGCCCAATTCAGGACACTTCTGAGTACATCAATTTCA 6580
b CTCTCTCAGACACCACTCTCTTGGAGCCGTTTCCAGGAGAGTCTCGAGTACATCAATTTCA 6879
y TGTCTATCTCTGTTGGCACTGATGAAGAACCTTTACAGTTCAGGGTTCCTGGAATTTCTACC 6640
b TGCCAAACCAAGTTGGCACCCAGCAGGAAGGCCCTTACAGTTCCTGGAATTTCTACC 6939
y AGTGCCACTCTGACAGGCTCACCAGAGTGCACCTCAACATCATATAGTGGAGCACTG 6700
b AGTGCCACTCTGACTGGCTTACCAGAGGGGTCACTCAACATCATATAGTGGAGCACTG 6999
y AAAGACACAGCAGAGGCATTAAGTTTCGGGAGAGGTTGTTTACCCTGGGCAACTCTCTCAAC 6760
b CAGAACAGAGGAGGCACAAAGTTTCGGGAAGGTTGTGACCTGGGCAACGCTGTGAGC 7059
y GAAGGCTTGAAACCACTACAGGATGACTCGTGTGTTGACCCCTACACAGTGTCCCATTTAT 6820
b GAAGGCTGAACAGGCTACAGATGACTCATGCTTTGACCCCTTACACGGTTTCCCATTTAC 7119
y GCGGTTGGAGATCAGTGGGACCAATGTTCTGAATCAGGCTTTAAACTGTTGTGCCAGTGC 6880
b GCCATTGGAGAGAGTGGGACGGTTGTCTGACGCTGGCTTTAAGCTCACATGCCAGTGC 7179
y TTAGGCTTTGGAAGTGGTCAATTTTCAGATGATGATTCATCTAGATGTTGCCATGACATGGT 6940
b TTGGGCTTTGGCAGTGGTCAATTTTCAGATGCGATTCATCTAAATGTTGCCATGACACGGT 7239
y GTCAATACAGAGATTGAGAGAGTGGGACCGTCAGGAGAGAAATGCCAGATGATGAGC 7000
b GTCAATACAGAGATTGAGAGAGTGGGACCGTCAGGAGAGAAATGCCAGATGATGAGC 7299
y TGACATGCTTTGGGAACCGAAAGAGAAATTCAGTGTGACCTCATGAGGCAAGTGT 7060
b TGACATGCTTTGGGAACCGAAAGAGAAATTCAGTGTGATGATGATGATGATGATGATGATG 7359
y TAGATGATGGGAGACATACAGTGGAGAGACAGTGGGAG -AAGGAATATCTCGGTGC 7119
b TAGATGATGGGAGACATACAGTGGAGAGACAGTGGGAG -AAGGAATATCTCGGTGC 7419
y CA-TTTGCTCTCTGCATGCTTTTGGAGCCAGCGGGCTGGCGCTGTGACAACTGCCGCA 7178
b CACTTTGCTCTCTGCATGCTTTTGGAGCCAGCGGGCTGGCGCTGTGACAACTGCCGCA 7479
y GACTGGG---GGTGAACCCAGTCCCGAAGCCACTACTGGCCAGTCTTACACCATGATT 7235
b GACTGGGCTGTGAACCCAGTCCCGATGGACCCAGCCGACCACTACACCATGATA 7539
y CTCAGAGATACCATCAGAGAACAACTAATGTTAATTTCCCAATTTGAGTGTTCATGC 7295
b CACAGAGATACCATCAGAGAACAACTAATGTTAATTTCCCAATTTGAGTGTTCATGC 7599
y CTTTAGATGTACAGGCTGACAGAGAAATTTCCCGAGAGTAAATCATCTTTCCAAATCAGA 7355
b CGCTAGATGTGCAAGCTGACAGAGACAGATTTCTCGAGAGTA---ATCTTTCCAGCCCCAC 7655
y GGAACAAGCATGCTCTCTGCGCAAGATCCATCTAACTGGAGTGAATTTAGCAGAC---C 7412
b CTTAAG--TGTCTCTCTACCAAGTTCATCCACCCAGTGTGATGTAGCAGACCTC 7713
y CAGTTAGAGTCTCTTTCTTTCTTTAAGCCCTTTGCTCTGGAGGAAGTGTCTCAGCTTC 7472
b CATCTTTAGTGGTCTTTTCAACCTTTAAGCCCTTTTGTCTCTGGAGCCATGTTCTCAGCTTC 7773
y AGCTCACTCAGCTTCTCCACGATCACCTCTG--GGAGTTTCTGAGGGTCTTTCTCAT 7531
b AGCAATTTAAGCTTCTCCAGAGTCCGCTGCTGGAGTGTGTGAGATCTTCTCTCTC 7833
y AATGAGGCTGCATTCAGCTGTTCTGCTTGGAAAGATTTCAATACCGCTCAGTATTT-- 7589
```

```
Db 7834 AATGGTGACAGTTGGTCACCCCTGTTCTGCTTTCAGGGTTTCAGTACTGCTCAGTGTGTTT 7893
Qy 7590 ----AAATGAAGTGAATCTTAAGATTTTGGATTCATATAGGAAAGCATATGACCCAA 7645
Db 7894 AAGAGAATCAAAAGTTCTTATGTTTGGTCTGGATCAATAGGAAACACAGGTAGCCAA 7953
Qy 7646 CCAAGATGCAAAATGTTTGAATGATATGACCAAAATTTTAAGTAGGAAGTCAACCCAA 7705
Db 7954 CTAGGA-GGAATGTACTGAATGCTAGTACCCAGAGCTTGGACGAAAGTCAACCCAGA 8012
Qy 7706 CACTTCTGCTTTCACCTTAAGTGTCTGGCCGCAATACTGTAGGAACAAG 7754
Db 8013 CACTCTGCTTTC-TTTTGCAATCTGACCTGACGACTGTGAGGACATG 8060

RESULT 3
AF130095 Homo sapiens clone FLC0562 PRO2841 mRNA linear HTC 08-MAY-2001
LOCUS
DEFINITION
ACCESSION AF130095
VERSION AF130095.1 GI:11493493
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2067)
Zhang, C., Yu, Y., Zhang, S., Wei, H., Bi, J., Zhou, G., Dong, C., Zai, Y.,
Xu, W., Gao, F., Liu, M. and He, F.
Functional prediction of the coding sequences of 75 new genes
deduced by analysis of cDNA clones from human fetal liver
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 2067)
Zhang, C., Yu, Y., Zhang, S., Wei, H., Bi, J., Zhou, G., Dong, C., Zai, Y.,
Xu, W., Gao, F., Liu, M. and He, F.
Direct Submission
Submitted (23-FEB-1999) Department of Experimental Hematology,
Institute of Radiation Medicine, Beijing Raiping Road 27, Beijing,
Beijing 100850, P. R. China
LOCATION/Qualifiers
FEATURES
source
1. 2067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FLC0562"
/tissue_type="liver"
/dev_stage="fetus"
872..1363
/note="predicted protein of HQ2841"
/codon_start=1
/evidence=not experimental
/product="PRO2841"
/protein_id="AAG35520.1"
/db_xref="GI:11493494"
/translation="MSESGFKLQCCLGFGSHFRCDSSRWCHNDNWNVYKIGERKWRDQ
RGWQMSKCTCLNGKGFCKDPHEATCYDDGKTYHYGEQWKEYLGAICSTCFGGQ
GEMWCDNCRPGGPPSPGTTGQSYNQYSQRVHQTNTNVNCPICEFPLDQADRED
SR"

Query Match 17.2%; Score 1342.2; DB 11; Length 2067;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 8; Indels 267; Gaps 1;

Qy 5897 ACTGACTACAGATCTACCTGTACACCTTGAATGACATGCTCGAGCTCCCTGTGTGTC 5956
Db 191 ATTAAATTAAAAATCTACTGTACACCTTGAATGACATGCTCGAGCTCCCTGTGTGTC 250
Qy 5957 ATCGACGCTCCACTGCCATTTGATGACCATCCAACTCCGCTTTCTTGGCCACACACCC 6016
Db 251 ATCGACGCTCCACTGCCATTTGATGACCATCCAACTCCGCTTTCTTGGCCACACACCC 310
```

Db 4540 AGGTCTCGATTCCCAACTGGTTTGTATCTCTGTATACACCGCAACTCAITTCAGTGT 4599  
Qy 4396 GCATCTGATTTGCTCTCTGAGCACCACATCACTGGTACAGATCGGCCATCATCCGAGCA 4455  
Db 4600 CCATGGGTGGCTCTCTCGGGCCCCCATCACCGGTATCATCATCGGCATCACGCCGAGCA 4659  
Qy 4456 CTTTCAGTGGGAGACCTCGAGAAGATCGGGTGGCCCCACTCTCGAAATTCACATCCCTCAC 4515  
Db 4660 TTCTGTCCGAAGACCCAGGAGGATCGAGTGGCCCTCGGGAAITTCATCACCTCTAC 4719  
Qy 4516 CAACCTCACTCCAGGCACAGAGTATGGGTACAGATCGTTGCTCTTAATGGCAGAGGGA 4575  
Db 4720 CAACCTTAATCCGGGCACCGAGTACGTGTTCAGCATCAITTCCTGTAAATGGCAGAGGA 4779  
Qy 4576 AAGTCCCTTATTTGATGGCCCAACATCAACAGATTTCTGTATGTTCCGAGGACACTGGAAGT 4635  
Db 4780 GAGCCCGCCTCATGTTGGCCAGCAGCAAGTTCCTGTATATTCAGAGATCTGGAGGT 4839  
Qy 4636 TGTGTCTGCGACCCACCACGCTACTGATCAGCTGGGATGCTCTCTGCTGTACAGTGA 4695  
Db 4840 CATTCGCTTCCACCCGCCACGAGCTGCTCATCAGTTGGGAACCCCTCTCGGTCTCTGTGG 4899  
Qy 4696 ATATTACAGGATCACTTACGAGAAACAGAGGAAATAGCCCTGTCCAGAGTTCACTGT 4755  
Db 4900 CTATTACAGAAATCACTACGAGAGACAGGAGAAATAGCCCTGTCCAGAGTTCACTGT 4959  
Qy 4756 GCCTGGGAGCAAGTCTACAGTACCATCAGCGGCTTAAATCCTGGAGTTGATPACCA 4814  
Db 4960 GCCCGAAGCAAGTCCACAGGCCATCAACAACTTAAACCCAGGAGCAGACTACACCA 5019  
Qy 4815 TCATGTGTATGCTGTACTGGCGGTGGAGACAGCCCGCAGCAGCAAGCCCAATTTCCA 4874  
Db 5020 TCACCTGTATGCTGTACTGGCGGTGGGACAGTCCAGCAGCAGCAAGCCAGTTTCCA 5079  
Qy 4875 TTAATTACCGAAACAGAAATGACAAACCATCCAGATGCAAACTGCAAGTGCAGGATTTCA 4934  
Db 5080 TCAATTATAAAGCAAAATGACAAGCGCTCCAGATGCAAGTGACAGATGTCAGGACA 5139  
Qy 4935 ACAGATTTAGTGTCAAGTGGCTGCTTCAAGTTCCTCTGTTACTGGTTACAGATPACCA 4994  
Db 5140 ACAGCATCAGTGTCAAGTGGCTGCTTCAAGTTCCTCTGTTACAGGCTACAGATGACCA 5199  
Qy 4995 CCACCTCCCAAAATGGACAGGACCAACAAATATAAACTGCAAGTTCAGATCAAAACAG 5054  
Db 5200 CCACCTCCCAAAATGGCTAGGACCATCAAAATATAAACTGCCAGTCCAGATCAAAACAG 5259  
Qy 5055 AAATGACTATTGAAGCTTGCAGCTTGCAGCCACAGTGGAGTATGTTAGTGTCTATGCTCAGA 5114  
Db 5260 AAATGACCATTCAGAGTTTGCACCCACTGTGGAGTACGTGGTTAGTGTATTATGCTCAGA 5319  
Qy 5115 ATCCAGCGGAGAGTCAAGCTCTGTTTCAGCTGCAAGTAAACCAATTCATCGCCCTA 5174  
Db 5320 ACCGAAACGGAAGAACCGCCCTGGTTTCAAACTGCAAGTACCAACATTCATGCGCCTA 5379  
Qy 5175 AAGGACTGCAATTCATGATGAGATGTCGATTCATCAAAATTCCTTGGGAAGCCAC 5234  
Db 5380 AAGGACTGCAATTCATGATGAGATGTCGATTCATCAAAATTCCTTGGGAAGCCAC 5439  
Qy 5235 AGGGCAAGTTTCCAGTACAGGTGACCTACTCCAGCCCTGAGATGCAATTCATGAGC 5294  
Db 5440 AGGGCAAGTTTCCAGTACAGGTGACCTACTCCAGCCCTGAGATGCAATTCATGAGC 5499  
Qy 5295 TATTCCTCTGACCTGATGTTGAAGAGACACTGCGAGCTGCAAGCTCTGAGCCGTTT 5354  
Db 5500 TTTTCTCTGACCTGATGTTGAAGAGACACTGCGAGAGCTGCGAGGCTCTGAGCCGGGT 5559  
Qy 5355 CTGAGTACAGTCAAGTGTGGTGGCTTGCAGATGATATGAGAGCCAGCCCTGATTG 5414  
Db 5560 CTGAGTACAGTCAAGTGTGGTGGCTTGCAGATGATATGAGAGCCAGCCCTGATTG 5619  
Qy 5415 GAACCCAGTCCACAGCTATTCTCTGCAACCACTGACCTGAAGTTTCACTCAGTCAACCCA 5474  
Db 5620 GAATCCAGTCCACAGCCTATTCTCTGGCCCCCAACCACTGAAGTTTCACTCAGTCAACCCA 5679

Qy 5475 CAAGCTCGAGCGCCAGTGGACACACCCAAATGTTTCAGCTCACTGGATATCGAGTGGCGG 5534  
Db 5680 CAGGTTTACTTGGCCAGTGGATAGACACCAAGTGTTCAGCTCACTGGCTACCGGTGGCGG 5739  
Qy 5535 TGACCCCAAGAGAGACCGGACCAATGAAGAAATCAACTTGGCTCTGACAGCTCAT 5594  
Db 5740 TGAACCCGAAGAGAGACAGGACCAATGAAGAAATCAACTTCTCCAGACAGCTCAT 5799  
Qy 5595 CGTGTGTATCAGACTTATGTTGGCCACCAATATGAAGTGAAGTGTCTATGCTCTTA 5654  
Db 5800 CGGTATGTTGTAGACTCATCTGGTGGCCATTAATACGAAGTCAAGTGTCTATGCTCTCA 5859  
Qy 5655 AGGACATTTGACAAAGCAGACAGCTCAGGAGTGTCACTCTGAGAGAAATGTCAGCC 5714  
Db 5860 AGGACACACTGACAAAGCAGACAGCCAGCCAGGAGTCACTACTCTGAGAAATGTTAGCC 5919  
Qy 5715 CACCAAGAGGCTCGTGTGACAGATGTACTGAGACCAACCATCACTTACCTAGCTGAGAA 5774  
Db 5920 CTTCAAGAAAGGCCCTGTGACGGAGCTTACAGAGACCAACCATCACTATTAGCTGAGAA 5979  
Qy 5775 CCAAGACTGAGACGATCACTGGCTTCCAAAGTGTGCGTTCAGCCAAATGGCCAGACTC 5834  
Db 5980 CAAAGACAGAGCAATCACTGGCTTCCAAAGTGTGCGCATCCAGCCAAATGGCCAGACTC 6039  
Qy 5835 CAATCCAGAAACCATCAAGCCAGATGTACAGAGCTACACCATCACTGCTTACCAACAG 5894  
Db 6040 CAGTTTCAAGAGGACTCAGCCCGGATTTAGAGCTTACACCATTAAGGTTTACAGCCAG 6099  
Qy 5895 GCATGACTTACAAAGTCTACCTGTATCACTTGAATGACAATGCTCGGAGCTCCCTGTGG 5954  
Db 6100 GCATGACTTACAAAGTCTACCTGTATCACTTGAACGAATGCCCGAGCTCGCCGTGA 6159  
Qy 5955 TATCGACCCCTTCCACTGCACTGATGACCATTCACCTGCTGCTGCTTCTGCGCCAGACAC 6014  
Db 6160 TCACTGATCCCTCCACCGCCATCGACGACCATTCACCTGCTGCTGCTTCTGAGCCAGAC 6219  
Qy 6015 CCAATTCCTTGTGTGTATCATGGCAGCCGCGCAGCTGCGCAGGATTAACGGCTACATCAT 6074  
Db 6220 CCAATTCCTTGTGTGTATCATGGCAGCCGCGCAGCTGCGCAGGATTAACGGCTACATCAT 6279  
Qy 6075 AGTATGAGAGCTGGTGTCTCTCCAGAGAAAGTGGTCCCTGCGCCCGCTGCTGCTCA 6134  
Db 6280 AGTATGAGAGCTGGTGTCTCTCCAGAGAAAGTGGTCCCTGCGCCCGCTGCTGCTCA 6339  
Qy 6135 CAGAGCTACTTACTGCTGCTGGA-ACCGGAAACCGAATATACAAATTTATGTCATTGCTCC 6193  
Db 6340 CCGAGGCCCACTTACTGCTTGGAGGCCAGGAAACCGAGTACACCATCTACGTCATTGCTCC 6399  
Qy 6194 CTGAAGAAATATCAGAAAGAGAGCCCTGATTTGGAAGGAAAGACAGACAGCTTCC 6253  
Db 6400 CTGAAGAAACAAATCAGAAAGAGTGAAGCCCTGATTTGGAGGAAAGACAGATGAGCTTCC 6459  
Qy 6254 CAACTGTAACCTTCCACACCCCAATCTTTCATGAGACAGAGATCTTGGATGTTCTTCC 6313  
Db 6460 CAACTGTTACCTTCCACACCCCAATCTTTCATGAGACAGAGATCTTGGATGTTCTTCC 6519  
Qy 6314 ACAGTTTCAAAAGACCCCTTTCATCCACCCCTGAGTGAACCTGGAATGATGATGATGAT 6373  
Db 6520 ACAGTTTCAAAAGACCCCTTTCATCCACCCCTGAGTGAACCTGGAATGATGATGATGAT 6579  
Qy 6374 GTTCTTGGCACTTCTGGTCAAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGAAAT 6433  
Db 6580 GTTCTTGGCAACCAACCCAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGAAAT 6639  
Qy 6434 GGTTTTGGGCGACCAACCGCCCAACCGCCCAACCGCCCAACCGCCCAACCGCCCAACCG 6493  
Db 6640 GGCTTTAGGCGGACCAACCGCCCAACCGCCCAACCGCCCAACCGCCCAACCGCCCAACCG 6699  
Qy 6494 TACCCCGGAGTGT----- 6507  
Db 6700 TACCTGCCGAATGTAGATGAGGAGGTCCTAAATCGGTGTCATGTTCCCAAGGGAGACGTGGAC 6759



2455 CGAATACAGCCAGTAGCTTTGGTCTCCCTGGTCTCAGCTTCGACACCGTGTCCGG 2514  
2380 TGAATACAGCCAGCAGCTTTGGTCTCCTGGTCTCAGCTTCGACACCGTGTCCGG 2439  
2515 ATTCCGGGTGGAAATAGCTAGTGAAGGAGGAGATGAGCAGCAGTACCTGGATCTTCC 2574  
2440 CTTCCGGGTGGAGTAGAGCTGAGCGAGGAGGAGATGAACACAGTACCTTGAICTCC 2499  
2575 AAGCAGCCACTTCTGTGAACATCCCTGAGCT- GCTTCTGGCCGAAATACATGTAA 2633  
2500 AAGCAGCCACTTCCGTGAACATTCCTGACCTGGCTCCCGGCGAAGATACATGTCA 2559  
2634 ATCTCTATCAGATATCTGAGGATGGGGAGCAGAGTTGATCTCTGCTACTTCACAAACA 2693  
2560 ATCTCTATCAGATATCTGAGAGGGGAACAGAGCTTGATCTGCTACTTCACAGACTA 2619  
2694 CAGCGCTGATGCCCTCTCTGACCCGACTGTGGAACAGTTGATGACACCTCAATGTGTG 2753  
2620 CAGCAGCTGACGCTCTCCAGACCTTACCCTGAGCAGGTTGATGATCTTCATCTGTG 2679  
2754 TTCGCTGGAGCAGACCCAGGCTCCCATCACAGGTACAGAAATAGTCTATTGCGCCATCAG 2813  
2680 TTCGCTGGAGTAGACCCAGGCACTTATCACAGGTATAGAAATGTCTATTCACTTCAG 2739  
2814 TAGAAGTAGACAGCAGAACTCAACTCTCTGAAACTGCAAACTCCGTCACCTCAGTG 2873  
2740 TAGAAGCAGTAGCAGAGCTCAACTCTCTGAAACGGCCAACTCCGTCACCTCAGCG 2799  
2874 ACTTGCACCTGCTGTTCAGTATAACATCACTATCTATCTGTGGAAGAAATCAAGAAA 2933  
2800 ACTTGCACCCGCTGTTCAGTATAACATCACTATCTATCTGTGGAAGAGAACAGGAGA 2859  
2934 GTACACCTGTTGTCATTCAACAAGAAACCACTGGCA CCCCAGCTCAGATACAGTGCCCT 2993  
2860 GCACCCGTTTTCATCCAAACAGAGACCACTGGCACCCCAAGATCTGATAACGTCCTCC 2919  
2994 CTCAGGAGCTTCAGTTTGTGGAG- TGACAGAGTGAAGTCAACATCATGTGGACA 3052  
2920 CTCGAGGACCTACAGTTTGTGGAAGTCACTGATGTGAAAGTCAACATCATGTGGACC 2979  
3053 CCGCTCAGAGTGCAGTGACCGGCTACCGTGTGGATGTGATCCCGTCAACCTGCTGCG 3112  
2980 CCTCTGATAGTGTGGTGTCTGGATACCGTGTGGAGTCTGCTGCTGACGCTGCCCCGG 3039  
3113 GAGCAGCGGAGAGGCTGCCATCAGCAGAAACACTTTCAGAACTCAACCGGCTGCTCC 3172  
3040 GAAATGGGAGAGGCTGCTGTCAACAGAAATACCTTTGCTGAAATCACTGGGCTGTCC 3099  
3173 CTTGGGCTACCTATTACTTCAAGTCTTTGCACTG- AGCCATGGGAGGAGAGCAAGCC 3231  
3100 CTTGGGCTCAGTACTCTTCAAGTCTTTGCTGTGCGAGCCAGGCGAGGAAAGCAATCC 3159  
3232 TCTGATGCTCAACAGAACCAACTGAGTGTCTCCCACTAACTCCAGTTTGTCAATGA 3291  
3160 TCTGACGGCAACAGAACCAACTGAGCTGCCACTAACTCACTCACTGAGTTGTCAATGA 3219  
3292 AACTGATTTACTGCTGCTGTGATGAGTCACTCCAGCTCGGCGCCAGATACAGGATACCG 3351  
3220 AACTGACAGAACAGTTCTGTGTAACGTGATCTCACTTCGAGCCCGTATAGCAGCTACCG 3279  
3352 ACTGACGCTGGCTTACCCGAGAGAGAGCAGCCAGCGAGTACATGATGGGTCCTCTGT 3411  
3280 ACTGACCGCGGCTGACCCGAGGAGGCGCAGCCCAAGCAGTACAAATGTGGGACCTTGGC 3339  
3412 CTCGAGTACCACTGAGGATCTGAGCCTGTGATCTGAGTACACCGTATCCCTCGTGGC 3471  
3340 CTCGAGTATCCCTGAGAAATCTGAGCCTGGTCTGAGTACACCTGACCTTGGTGGC 3399  
3472 CATAAAGGCGCAACAGAGAGCCCAAGCACTGAGGCTTTTACACACTGAGCCTGG 3531  
3400 TGTGAAGGGAACAGCAGAGTCCCAAGCCACCGGAGTCTTTAGTACCTCGCCTCT 3459  
3532 GAGCTCTATTCCACTTACACACCGGAGTCACTGAGACCACTTGTGATCAGATGGAC 3591

3460 GCGTCCATTCCACCTTTACACACCGAGGTGACAGACCAATTTGTATCACCCTGGAC 3519  
3592 GCGTCTCAAGAATTGGTTTAAAGCTGGTGTACGACCAAGCCAGGAGGAGGACCC 3651  
3520 CCGGCTCCAAGGATTTGGCTTCAAGCTGGTGTACGACCGAGCCAGGAGGTGAGGACCC 3579  
3652 ACAGAAAGTCACTTCAGACTCAGAAAGCATCGTTGTGTCCGGCTTGACTCCAGGAGTGA 3711  
3580 CCGAAGTCACTTCAGACTCTGGAGCATTTGTGTCTGCTTGCATCCAGCGGTGA 3639  
3712 ATAGCTCTACCAATCCAGTCTGAGAGATGACAGGAAAGAGATGCCCAATTTGAAA 3771  
3640 ATACACTTACACCATCCAGTCTGAGATGCGCAGGAGAGATGACCCGATTTGCAA 3699  
3772 CAAAGTGGTCAACCACTGTCTCCACCAACAACTTGCACTCTGGAGCAACCCCTGACAC 3831  
3700 CAGATAGTGAACCCGCTGTCTCCACCGACCACTTGCACTCTGGAGCAACCCCTGACAC 3759  
3832 TGGAGTCTCAGAGTCTCTGGAGAGGAGCAGCAGCCAGACATTTACTGGTTATAGAA 3891  
3760 TGGAGTCTTACTGTCTCTGGAGAGGAGCACTACCCAGATATCACTGGCTACAGAA 3819  
3892 TACCACAACCCCTACAAACCGCCAGCAGGCAAAATTTCTTGAAGAAAGTGGTCCATGCTGA 3951  
3820 AACTACTACCCCAAGAACCGGAGCAGGAGGAGCCTCTCTGGAAGAGTGTCTCCATGCTGA 3879  
3952 TCAGAGTCTCTGCACTTTTGAATAACCTGAGTCCCGGCTCGAGTACATGTCAGTGTTTA 4011  
3880 TCAGAGTCTCTGCACTTTTGAATAACCTGAGTCTCTGGCTCGAGTACAACTGAGTGTTA 3939  
4012 CACTGTCAAGGATGACAGGAAAGTGTCCCTATCTCTGATACCATCATCCCA----- 4063  
3940 CACTGTCAAGATGACAGGAAAGTGTCCCTATCTCTGATACCGTTGTCCAGAGGTGCC 3999  
4064----- 4063  
4000 CCAGTCACTGACCTAAGCTTTGTGTGATATAAATGATTCAAGCATCGGCTGAGGTGGAC 4059  
4064----- 4063  
4060 CCGCTAACTCTTCCACCACTTATCGGGTACCGAATCACAGTAGTTGGGCGAGGAGG 4119  
4064----- 4063  
4120 GATCCCTATTTTGAAGATTTTGTGGACTCTCAGTAGGATACACAGATTACAGGGG 4179  
4064----- 4063  
4180 CTGGAGCCTGSCATTTGACATGACATCAGCGTTATCACTCTCATTAATGGCGAGAGAT 4239  
4064----- GCTGTTCTCTCTCCACTGCTCCACTGACCTGCGATTCAAC 4096  
4240 GCCCTACTACACTGACACAGCAAAACGGCTGTCTCTCCACGGATCTGCGATTCAAC 4299  
4097 AACATTGTCTCAGACACCATCGTGTACCTGGGCTCCACCC-CCATCCATTGATTAAAC 4155  
4300 AATATCGGTCCAGACACGATCGGGTCACTTGGGCGCCGCTCCCTCCATCGAGTGAC 4359  
4156 CAACTTCTGGTGGCTTACTCACTGTGAAAATGAGGAAAGATGTGAGAGTTGTCAAT 4215  
4360 CAACTCTTGGTGGCTTACTCACTGTGAAAGATGAGAGGAGCTTGCAGAGCTATCCAT 4419  
4216 TTCTCTTTCAGCAATGAGTGGTCTTAACTCTCTCTCTGCTGGTACAGAAATGTAGT 4275  
4420 TTCACTTTCAGCAATGCGGTGCTTAACTCTCTCTCTGCTGGGACAGAAATCTTAGT 4479  
4276 GAGTGTCTCCAGTGTCTACGAAACACATGAGAGCACCTCTTAGAGGAAACAGAAAC 4335  
4480 CAGTGTCTCCAGTGTCTACGAAACACATGAGAGCATCCCTCTCCGGGGAAGGAGAAAC 4539  
4336 AGGTCTTGTATCCCACTGGCATTTCTTGTGATTAATCTGCAACTCTTTTACTGT 4395

Db 168 GCGACTCTCTCCATCCACTCAAGATGCTCAGGGGTCCGGGACCCGGCGCTGCTGC 227  
 Qy 300 TGCTGGCCGCTCTGTGCTGGGACAGCGGTGCCCTCCACGGGAGCCTCGAAGAGCAAGA 359  
 Db 228 TGCTGGCAGTCTGTGCTCTGGGACCTCGGTGCGCTGCACCGAGCGGAGAGCAAGA 287  
 Qy 360 GCGAGGCTCAGCAATGGTGTGAGCCCGAGTCCCGGTGGCTGTGAGTCAAGAGCAGCCG 419  
 Db 288 GCGAGGCTCAGCAATCGTGGAGCTCAATCCCGGTGGCTGTGAGTCAAGAGCAGCCG 347  
 Qy 420 GTGTGTATGACAAATGGAACACATATCATAGATAAATCAACAGTGGAGCGGACCTACCTAG 479  
 Db 348 GCTGTTTTGACATGGAAGCACTATCAGATATATCAGCATGCGAGCGGACCTACCTAG 407  
 Qy 480 GCAATGCGTGTGTTGTTACTTGTATGAGAGAGCCGAGGTTTTAACTGCGAAGATTAAC 539  
 Db 408 GCAACGCCCTGG--TGTAACCTGCTATGAGGAAGCCGGGTTTTAACTGCGAGAGCAAGC 465  
 Qy 540 CTGAGGCTGAGAGACTTGTCTTGTGACAACTGACCTGAGGAGCACTTACCGAGTGGTGACA 599  
 Db 466 CTGAGCCTGAAGAGACTTGTCTTGTGACAACTGACCTGAGGAGCACTTACAAAGTGGTGACA 525  
 Qy 600 CTTATGAGCGTCTTAAAGACTTCATGATCTGGGACTGTACCTGCAATCGGGCTGGGCGAG 659  
 Db 526 CTTATGAGCGCTTAAAGATTCCATGATCTGGGACTGTACCTGCAATCGGGCTGGGAGAG 585  
 Qy 660 GGAGATAGCTGTACCATCGCAACCGCTGCCATGAGGGGCTCAGTCTACAGATTG 719  
 Db 586 GCAGATCAGCTGTACCATGTGCAATCGCTGCCATGAAAGGGGCTCAGTCTACAGATTG 645  
 Qy 720 GTGACACCTGGAGGAGACCACTGAGCTGTGTTGATGTTAGAGTGTGTGTCTTG 779  
 Db 646 GCGACAGTGGAGGAGGCCACATGAGACTGTGTGGCTACATGTTAGAGTGTGTGTCTTG 705  
 Qy 780 GTAATGGAAGAGAGATGAGCTCTGCAAGCCCATAGCTGAGAGTGTGTTGATCATGCTG 839  
 Db 706 GAATGGAAGAGGAGATGAGCTCTGCAAGCCCATAGCTGAGAGTGTGTTGATCATGCTG 765  
 Qy 840 CTGGGACTTCTTATGTGTGCGAGAAACGTGGGAGAACCTTACCAAGGCTGGATGATGG 899  
 Db 766 CTGGGAGTCTTACGTCTGTGGGAGACCTGGGAGAACCTTACCAAGGCTGGATGATGG 825  
 Qy 900 TAGATTGTTACTTGTCTGGAGAGGAGCGGACCGATCATCTTGCACTTCTAGAAATAGAT 959  
 Db 826 TGGACTGTACTTGTCTAGGAGAGGCAATGGAGCGATCATCTGTACTTCCAGAAACAGAT 885  
 Qy 960 GCAACGATCAGGACACCAAGGACATCTATAGAAATGGAGACCTTGGAGCAAGAGGATA 1019  
 Db 886 GCAACGATCAGGACACCAAGGACATCTATAGAAATGGAGACCTTGGAGCAAGAGGATA 945  
 Qy 1020 ATCGAGGAAACCTGTCTCAGTGCACTGCAAGGCAACCGGCGAGAGAGTGGAAAGTGTG 1079  
 Db 946 ACCGAGGAAACCTGTCTCAGTGCACTGCAAGGCAATGGAGAGGAGTGGAAAGTGTG 1005  
 Qy 1080 AGAGGACACCTGTGTGAGACACATCGAGCGGATCTGGCCCTTTCACCGATGTTCTGTG 1139  
 Db 1006 AGGACACCTGTGTGAGAGTGTCTAGCGGATCTGGCTCTTCTACTGATGTCCGAA 1062  
 Qy 1140 CAGCTGTTTACCAACCGCAGCTTCAACCCAGCTCTCTCTTATGAGGCACTGTGTCAAG 1199  
 Db 1063 CAGCTATTTACCAACCGCAGCTTCAACCCAGCTCTCTCTTACGAGGCACTGTGTCAAG 1122  
 Qy 1200 ACAGTGTGTGTCTACTTGTGGGATGAGTGTGAGGACCAAGGAAATAG--CAA 1258  
 Db 1123 ACAGTGTGTGTCTACTTGTGGGAAATGAGTGTGAGTGTGAGTGTGAGGAAATAGCCAA 1182  
 Qy 1259 ATGCTTTGACGCTGCTGGGCAACGGAGTGTGAGTGTGAGGACAGCTGTATACCGAGCT 1318  
 Db 1183 ATGCTGTGACAGTGTGCTGGGCAATGGCTGAGTGTGAGGACAGCTGTGAGGAGCT 1242  
 Qy 1319 TACGTTGGCACTCAATGAGAGGCC--ATGTGTCTTACCTTACCTCAATGGGAGGAC 1377  
 Db 1243 TATGGTGGCAATTCAAACGGGAGGCCCTGTGTCTCTCCGTTTCACTCAACGGTAGGAC 1302

Qy 1378 GTTCTACTCTCTGACACACAGAAAGGGGAGCAGGACGACATCTTTGGTGCAGACACTTC 1437  
 Db 1303 CTTCTATTCTCTCACCACCGAAGGGCGGCAAGACGACATCTGTGGTGTAGCACAATTC 1362  
 Qy 1438 GAATATATGAGCAGGACCAAGAAATCTCTTTCTGACAGACACACTGTTTGGTTGAGAC 1497  
 Db 1363 CAATACGAAACAGACCAAGAGTATCTCTTCTGACAGACCAATGCGGTTTGGTTGAGAC 1422  
 Qy 1498 TGAGAGAGAAATTCGAATGCTGCTGTGCGCACTTCCCTTCTCTATACAAACACCA 1557  
 Db 1423 TCAGGCGGAAATTCGAATGCTGCTGTGCGCACTTCCCTTCTCTATACAAACCGGAA 1482  
 Qy 1558 TTACACTGATGTCACCTTCTGAGGCGAGAGAGACAAACATGAAGTGTGTGGGACCA--CAC 1616  
 Db 1483 TTACACCGACTGTACTTCTGAGGCTGCGAGGCAACATGAATGTTGGTGGCACCACCC 1542  
 Qy 1617 AGAATCTATGATCCGACCAAGAGTTTGGTCTTCTGCCCCATGCTGCCACGAGGAAATCT 1676  
 Db 1543 AGAATCTACGATCCGATCAGAAATTTGGATTTCTGCCAATGCTGCCACGAGGAGATCT 1602  
 Qy 1677 GCAACCAATGAAGGGTCTGTTACCGCATTTGGAGATCAGTGGGATAAGCAGCATGACA 1736  
 Db 1603 GCAACCAATGAAGGGTCTGTTATCGCATTTGGGATCAGTGGGATAGCAGCATGACC 1662  
 Qy 1737 TGGGTCACTGATGAGTGTCACTGTGTTGGGAATGCTGTGGGGAATGGAATGCAATG 1796  
 Db 1663 TGGGCCACATGATGAGTGTCACTGTGTTGGGAAACGCTGTGGGAAATGGGCTGCACTCC 1722  
 Qy 1797 CCTA--CTGCGGCTTCCGAGATCAGTGTCTTGTGATGACATCACTTACATGTGAACGAC 1855  
 Db 1723 CCTAGTCTCCAGCTCCGAGACCACTGATGCTGTTGATGACATTAATTTCAATGTGAACGAC 1782  
 Qy 1856 ACATTTCCAAAGCGTCTGAAAGGGGCGCATGCTGAACTGTACATGCTGTGGTCTAGGGT 1915  
 Db 1783 ACCTTTCCAAAGCGTCAAGAGGAGGACATATGCTGAACTGTACCTGTGTGGTCTAGGGC 1842  
 Qy 1916 CCGGCGAGTGTGAGTGTGATCCGCTGCAACATGCGAGGATTCAGAGATTCAGAGTGGAGT 1975  
 Db 1843 CCGGCGAGATGGAAGTGTGACCCCATTTGACCAAGTGTGCAAGATTCAGAGACCCGGAAT 1902  
 Qy 1976 TATCAAAATTCGAGATTCATGGGAGAAATGTCATGCTGTGTGAGATACCAAGTCTACTGC 2035  
 Db 1903 TACAGATTTGCTGACTCTCTGGGAGAAATTTGTCATGTTGTCGATACCAAGTCTACTGC 1962  
 Qy 2036 TATGGCCGCTGG--CATTTGGGAGTGGCATTTGCCAACTTTACAGACCTTATCCAGCTCAAG 2094  
 Db 1963 TACGGCCGCTGGCCATCGGGAGTGGCATGTCAACTCTGCAAGCTTACCCAGGCAAC 2022  
 Qy 2095 TGTCTCTGTCGAAGTATTTATCACTGAGACTCCGAGTCAAGCCCACTCCACCCCATCCA 2154  
 Db 2023 TGGACCTGTCCAAGTAAATATCAAGGAGCCCCCAGCCAGCCCAATTTCCACCCCATCCA 2082  
 Qy 2155 GTGGAATGCAACACAGCCATCTCACAATTTCCAAAGTACATTTCTCAGGTGGAGACTTAAAA 2214  
 Db 2083 GTGGAATGCCCCGAGCCTTCAACATCAACAGTATCAATTTCTCAGATGGAGACTTAAAA 2142  
 Qy 2215 TTCTGTAGGCGGTGGAAGGAGCTACATACAGGCGCACTTAAACTCTCTACCACTCAATCAA 2274  
 Db 2143 CTCTACGGGTGCTGGAAGGAGCTACCAATCTGCGCACTTAACTCTCTATACCATCAA 2202  
 Qy 2275 AGGCTTGAAGCCTGCTGTGTATACGAGGCGCAGCTCATCAGCATCCAGCATCGGCCA 2334  
 Db 2203 AGGCTTGAAGCCTGCTGTGTATATGAGGAGAGCTCATCAGCATCCAGCATGAGGCCA 2262  
 Qy 2335 CCAAGAGTGAATGCTGCTTGTGACTTCAACAGGCGCAGCTTCAACAGGCGCAGCAGCTGTGACCA 2394  
 Db 2263 CAGAGAGTGAACGCTTTCGACTT---CAGACAGGCGCAGCAGCTGTGAGCAGCA 2319  
 Qy 2395 CACCGTGAAGGAGAGAGCTCCCTTTCTCTCTTGTGGCCACTTCTGAATCTGTGAC 2454  
 Db 2320 CACCGTGAAGGAGAGAGCTGGCCCTTCTCTCTTGTGGCCACTTCTGAATCTGTGAC 2379

```
b 7601 CTGATAGAGACGATTCTCGAGAGTA-----ATCTTTCCAGCCCCCTACAG--TGCT 7654
y 7371 CTCTGCCAGATCCATCAACTGAGTGATCTTAGCAGAC---CCAGCTTAGAGTTCTT 7427
b 7655 CTCTACCAAGGTCAATCCACCCAGCAGATGTTAGCAGACCTCCATCTTTAGTGGTC 7714
y 7428 CTTTCTTTTAAAGCCCTTGTCTGGAGGAAGTTCTCCAGTTCAGTCAACTCAGC 7487
b 7715 CTTTCAACCCTTAAGCCTTTGCTCTGGAGCCTATGTTCTCAGCTTCAGCACAATTTACAGC 7774
y 7488 TTCTCCAGCATCACCCCTG-GGAGTTTCTGAGGGTTTTCTCATAAATGAGGGTGCACA 7546
b 7775 TTCTCTAGCATGCCCCCTGGATGTTTGGAGTCTTCTCCCAATGGTGCAGTTGG 7834
y 7547 TTGCTCTGTTCTGCTTGAAGTATTCAATACCCCTCAGTATTTT-----AAATGAAGTGA 7600
b 7835 TCACCCCTGTTCTGCTTCAGGTTTCTAGTACTCTCAGTGTCTTGAAGAGATCAAAAGT 7894
y 7601 TTCTAGATTGTTGGTCAATAGGAAGCATATGAGCAGCAACCAAGATGCAAAAGT 7660
b 7895 TCTTATGGTTTGGTCTGGGATCAATAGGGAACACAGGTAGCACTAGGA-GGAAATGT 7953
y 7661 TTGGAATGATATGACCAAAATTTTAAAGTAGGAAGTACCCCAACACTTCTGCTTTCAC 7720
b 7954 ACTGAATGCTAGTACCAAGACCTTCAGCAGGAAGTCAACCCAGACACTCTGCTTTC-T 8012
y 7721 TTAAGTGTCTGGCCCAATAGTCTAGGGAACAAG 7754
b 8013 TTGGCATCTGACCTGCAGCAGTCTCAGGACATG 8046

RESULT 2
LOCUS AK090130 8329 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus 5 months female bone marrow stroma cell CRL-2028
SR-4987 cDNA, RIKEN full-length enriched library, clone:G431002M24
product:fibronectin 1, full insert sequence.
ACCESSION AK090130
VERSION AK090130.1 GI:26105744
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
JOURNAL High-efficiency full-length cDNA cloning
MEDLINE Meth. Enzymol. 303, 19-44 (1999)
PUBMED 98279253
REFERENCE 2
AUTHORS Carninci,P., Shibata,K., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itch,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagao,K., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itch,M.,
Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yanamoto,R., Masumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Chata,B., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
```

```

PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 8329)
ADACHI,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Inotani,K., Ishii,Y., Itch,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numata,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akaira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
source
1. 8329
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BDF1"
/db_xref="PANTOM,DB:G431002M24"
/db_xref="MGI:2429036"
/db_xref="taxon:10090"
/clone="G431002M24"
/sex="female"
/cell_line="CRL-2028 SR-4987"
/tissue_type="stroma cell"
/tissue_type="bone marrow"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="5 months"
misc_feature
194..7640
/note="fibronectin 1 (MGI:95566, GB|X93167, evidence:
BLASTN, 99%, match=4434)
putative"
polyA_signal
8310..8315
/note="putative"
polyA_site
8329
/note="putative"
ORIGIN
Query Match 66.5%; Score 5187.4; DB 11; Length 8329;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 6509; Conservative 0; Mismatches 1046; Indels 414; Gaps 28;
QY 184 GGAATTTTCCGAGTTGCCAAGGAAGCAAACTGGTGGCACTTCCTCCCGTGC 243
Db 108 GGACTTTTCCGAGCTGCGAGGAGCCACAGTTCGTGGCCACTTCCTCTGGGAG 167
QY 244 GCGGCTCTCC-CCACACGCTCTCAACATGCTTAGGGGTCGCGGCGCCG- --GGCTGCTGC 299
```







QY	1140	CAGCTGTTTACCAACCGCAGCTCACCCCGAGCCTCTCTCTCTGCGCACTGTGTACAG	1199	CGGTCGCTGGAAGGAAGCTACCATTCCTTGCCCACTTAACTCTATACCATCAAAGGCC	2204
Db	1165		1174		
QY	1200	ACAGTGGTGTGGTCTACTCTGTGGGATGCGATGGCTGGAAGACACAAAGGAATAAGCAAA	1259		
Db	1125	ACAGTGGTGTGGTCTACTCTGTGGGATGCGATGGCTGGAAGACACAAAGGAATAAGCAAA	1184		
QY	1260	TGCTTTGCAAGTGGCTGGCAACGGAGTCAGCTGCCAAGAGACAGCTGTAAACCCAGACTT	1319		
Db	1185	TGCTGTGCAAGTGGCTGGCAACGGAGTCAGCTGCCAAGAGACAGCTGTAAACCCAGACTT	1244		
QY	1320	ACGGTGGCAATCAAAATGGAGAGCCATGTGTCTTACCAATTCACCTAACAATGGCAGGACGT	1379		
Db	1245	ATGGTGGCAATTCAAAAGGGGAGCCCTGTGTCTCTCCGTTTCCACCTAACAAGTAGGACCT	1304		
QY	1380	TCTACTCTGACACACAGAAAGGGCGACAGGAGGACATCTTTGGTCAGACACACTTCGA	1439		
Db	1305	TCTATTCTGACACACAGAAAGGGCGGCAAGACGACATCTGTGTGTAGACAACTTCCA	1364		
QY	1440	ATTATGACGAGGACACAGAAATACTCTTTCTGCACAGACCACACTGTTTTGGTTCAGACTC	1499		
Db	1365	ATTACGACACAGACAGAAATATCTTCTGCACAGACCACACTGTTTTGGTTCAGACTC	1424		
QY	1500	GAGGAGGAAATTCGAATGGTGGCTGTGCCATCTTCCCTTCTATACAAACCAACAATT	1559		
Db	1425	GAGGCGGAAATTCGAATGGTGGCTGTGCCATCTTCCCTTCTATACAAACCAACAATT	1484		
QY	1560	ACACTGATTGCACTTCTGAGGCGAAGAGACAAACATGAAGTGTGTGGGACCAACACAGA	1619		
Db	1485	ACACCGACTGTACTTCTGAGGGTCGACGGGACAAACATGAATGTGTGGGACCAACCCAGA	1544		
QY	1620	ACTATGATCGCGACACAGAAATTTGGTTCTGCCCATGGCTGCCACGAGGAATCTGCA	1679		
Db	1545	ACTACGATCGCGATCAGAAATTTGGATTCTGCCCAATGGCTGCCACGAGAGATCTGCA	1604		
QY	1680	CAACCAATGAGGGGTCAATGACCCATTTGGAGATCAGTGGGATGAGCAGCATGACATGG	1739		
Db	1605	CAACCAATGAGGGGTCAATGACCCATTTGGAGATCAGTGGGATGAGCAGCATGACATGG	1664		
QY	1740	GTCACATGATGAGTGCAGCTGTGTGGGAAATGGTGTGGGGAATGGACATGCAATGGCT	1799		
Db	1665	GCACATGATGAGTGCAGCTGTGTGGGAAATGGTGTGGGGAATGGGCTGCAATGCCCT	1724		
QY	1800	ACTCCAGCTTCGAGATCAGTGCATTTGATGATGATGATGATGATGATGATGATGATGATG	1859		
Db	1725	ACTCCAGCTTCGAGATCAGTGCATTTGATGATGATGATGATGATGATGATGATGATGATG	1784		
QY	1860	TCCACAAGCGTTCATGAGAGGGGCGCATGCTGAACTGTATACATGCTTCGGTCAGGGTGGG	1919		
Db	1785	TCCACAAGCGTTCATGAGAGGGGCGCATGCTGAACTGTATACATGCTTCGGTCAGGGTGGG	1844		
QY	1920	GCAGTGGGAAGTGTGATCCCTCGACCAATGCGAGATTCAGAGATTCGAGATGGAGCTTTATC	1979		
Db	1845	GCAGATGGAAGTGTGATCCCTCGACCAATGCGAGATTCAGAGATTCGAGATGGAGCTTTATC	1904		
QY	1980	AAATTCGAGATTCATGGGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG	2039		
Db	1905	AGATTGGTGTACTCTGGGAGAAATTTGTGATGATGATGATGATGATGATGATGATGATGATG	1964		
QY	2040	GCGTGGCATTTGGGAGTGGCATGTCACCACTTTACAGACCTATCCAGCTCAAGTGGTC	2099		
Db	1965	GCGTGGCATTTGGGAGTGGCATGTCACCACTTTACAGACCTATCCAGCTCAAGTGGTC	2024		
QY	2100	CTGTCGAAGTATTTATCACTGAGACTCCGAGTCAGCCCAACTCCACCCCACTCCAGTGGGA	2159		
Db	2025	CTGTCGAAGTATTTATCACTGAGACTCCGAGTCAGCCCAACTCCACCCCACTCCAGTGGGA	2084		
QY	2160	ATGCAACACAGCCATCTCAATTTCAAGTACATTTCTCAGGTGGAGACCTAAATTTCTG	2219		
Db	2085	ATGCCCCGAGGCTTCAACATCACCAGTACATTTCTCAGATGGAGACCTAAATTTCTA	2144		
QY	2220	TAGGCCGTTGGAAGGAGCTTACATACCGGCCACTTAACTCTTACACCACTCAAAGGCC	2279		

Db	2145	CGGTCGCTGGAAGGAAGCTACCATTCCTTGCCCACTTAACTCTATACCATCAAAGGCC	2204		
QY	2280	TGAAGCCTGGTGTGTATACGAGGCGCAGCTCATCAGCATCAGCAGTACGCGCACCAAG	2339		
Db	2205	TGACCCAGGTGTGTATATGAGGACAGCTCATCAGCATCAGCAGTATGCGCACAGAG	2264		
QY	2340	AAGTGACTCGCTTTGACTTACCTACCAACACAGCACACCTGTGTACCAAGCAACACCG	2399		
Db	2265	AAGTGACACGCTTCGACTT---CACCAACAGCGCCAGCACCCCTGTGACCAACACCG	2321		
QY	2400	TCACAGGAGAGAGACTCCCTTTTCTCTCTTGCCCACTTCTGAACTCTGTGACCGAAA	2459		
Db	2322	TGACCCGAGAGACTCGGCCCTTACTCTCTGTGTGGCCACTTCTGAACTGTAACTGAAA	2381		
QY	2460	TCACAGCCAGTATGCTTTGTGGTCTCTCTGGGTCTCAGCTTCCGACACCGTGTCCGATTC	2519		
Db	2382	TCACAGCCAGCAGCTTTGTGTCTCATGGTCTCAGCCTCCGACCCGCTGTGAGGCTTCC	2441		
QY	2520	GGTGGGAATATCAGCTGAGTGGAGGAGAGATGAGCAGTACCTGGATCTTCCAGACA	2579		
Db	2442	GGTGGAGTATGAGCTGAGCGAGGAGGAGATGAACCAAGTACCTTGTCTCTCCCAAGCA	2501		
QY	2580	CAGCCACTTCTGTGAACATCCCTGACCTGCTTCTTGCGCGGAAAAATACATGTAAATGTCT	2639		
Db	2502	CGGCCACTTCCGTGAACATTCCTGACCTGCTCCGCGCAGAAAAGTACATTTGTCAATGTCT	2561		
QY	2640	ATCAGATATCTGAGGTGGGAGCAGAGTTTGTATCTGTCTACTTCAACAAACAAAGCCG	2699		
Db	2562	ATCAGATATCTGAGGAGGAAAAACAGAGCTTGTATCTGTCTACTCAAGACTACAGCAC	2621		
QY	2700	CTGATCCCTCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAATTTGTTTGGTCT	2759		
Db	2622	CTGACCTCTCTCAGACCTCTACGTTGACCAAGTTGATGATGATGATGATGATGATGATG	2681		
QY	2760	GGAGCAGACCCAGGCTCCCATCAGAGGTACAGATAGTCTTATTCGCCATCAGTAGAAG	2819		
Db	2682	GGAGTAGACCCAGGACCTATACAGGGTATAGAAATTTGTCTTATTCACCTTCACTAGAG	2741		
QY	2820	GTAGCAGCAGAACTCAACCTTCTCTGAAACTGCAAACTCCCTCACCTCAGTGACTTC	2879		
Db	2742	GCAGTAGCAGAGCTCAACCTCCCTGMAACGGCAACTCCCTCACCTCAGGACCTGC	2801		
QY	2880	AACTGGTGTTCAGTATTAACATCATCTATCTGTCTGTGGAAGAAATCAAGAAAGTACAC	2939		
Db	2802	AGCCCGTGTTCAGTATCAACATCATCTATCTGTCTGTGAGGAGAACCCAGGAGGACAC	2861		
QY	2940	CTGTGTCTATTCAACAGAAACCACTGGCAACCCAGCTCAGATACAGTCCCTCTCCCA	2999		
Db	2862	CCGTTTTCTATCCAAAGAGACCACTGGCAACCCCAAGATCTGATACGTCCTCCCTCCGA	2921		
QY	3000	GGGACCTGCAAGTTTGTGGAAGTACAGAGGTGAGGTCAACATCATGTGACACCCGCTG	3059		
Db	2922	CGGACCTTACAGTTTGTGGAAGTACAGAGGTGAGGTCAACATCATGTGACACCCCTCCG	2981		
QY	3060	AGAGTCAGTACCGGCTACCGGTGATGATCCCGTCAACCTGCGCTGGCGAGCAGC	3119		
Db	2982	ATAGTGGTGTCTGGATACCGTGTGAGGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG	3041		
QY	3120	GGCAGAGGCTGCCCATCAGCAGGAACACCTTTCGAGAGTCAACCGGCTGTCCCTGGGG	3179		
Db	3042	GGCAGAGGCTGCCCTGTCAACAGAAATACCTTTCTGTAATCACTGGGCTGTCCCTGGGG	3101		
QY	3180	TCACCTATTCTCAAGTCTTTGCGAGTACCTGAGGAGGAGCAAGCTCTTGAATG	3239		
Db	3102	TCAGTACCTCTTCAAGTCTTTGCGAGTACCTGAGGAGGAGCAAGCTCTTGAATG	3161		
QY	3240	CTCAACAGAACCAAACTGGATGCTCCCACTAACTTCCAGTTTGTCAATGAATGATTT	3299		
Db	3162	CACAACAGAACCAAACTCGAGCTCCCACTAACTTCCAGTTTGTCAATGAATGATTT	3221		
QY	3300	CTACTGTCTGGTGAATGACTCCACTCGGCCCAAGATAACAGGATACCGACTGACCG	3359		



REFERENCE  
AUTHORS  
3 Shibata,K., Itoh,M., Aizawa,K., Nagao,K., Sasaki,N., Carninci,P.,  
Konno,H., Akiyama,J., Nishi,K., Kitaura,T., Tashiro,H., Itoh,M.,  
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujiwara,S., Inoue,K., Togawa,Y., Iizawa,M., Ohara,E., Watabiki,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
TITLE  
RIKEN integrated sequence analysis (RISA) system--384-Format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
PUBMED  
REFERENCE  
AUTHORS  
4 The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
TITLE  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
TITLE  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 583-573 (2002)  
6 (bases 1 to 8315)  
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,  
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M.,  
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,  
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
Okazaki,Y., Saito,R., Saito,H., Sakai,C., Sakai,K., Sakazume,N.,  
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)  
COMMENT  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.go.jp/  
URL:http://fantom.gsc.riken.go.jp/  
FEATURES  
source  
1. .8315  
/organism="Mus musculus"  
/mol\_type="rRNA"  
/strain="BDF1"  
/db\_xref="FANTOM\_DB:G431004D19"  
/db\_xref="MG1:2429041"  
/db\_xref="taxon:10090"  
/clone="G431004D19"  
/sex="female"  
/cell\_line="CRL-2028 SR-4987"  
/cell\_type="stroma cell"  
/issue\_type="bone marrow"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="5 months"  
194. .7626  
misc\_feature  
/notes="fibronectin 1 (MGD|MG1:95566, GB|X93167, evidence:  
BLASTN, 99%, match=4434)  
putative"  
polya\_signal  
8296. .8301

polyA\_site  
8315  
/note="putative"  
ORIGIN  
Query Match 66.3%; Score 5324.8; DB 11; Length 8315;  
Best Local Similarity 81.8%; Pred. No. 0;  
Matches 6504; Conservative 0; Mismatches 1052; Indels 398; Gaps 15;  
184 GGATTTTTCGGCAGGTTGGAAAGGAAAGCAAACTTGGTGGCAACTTGGCTCCCGTGGG 243  
108 GGACTTTTTCGGCAGGTTGGAGGGGACCCACAGTTTCTGTGGCCACTTGGCTCTCTGGGAG 167  
244 GGCCTCTCTCC-CCCACTCTCTCAATGCTTAGGGTTCGGGGCCCG--GCCTGCTGC 299  
168 GGCACCTCTCTCCATCCACTCAAGATGCTCAGGGGTTCGGGACCCCGGCGCTGCTGC 227  
300 TGCTGGCGCTCTGTGCTCGGACACAGCGTGCCTCCACGGGAGCTCGAAGAGCAAGA 359  
228 TGCTGGCAGTCTGTGCTCGGACCTCGTGGCTGCACCGAAGCCGGGAAGAGCAAGA 287  
360 GGCAGGCTCAGCAATGGTTTCAGCCCCAGTCCCCGGTGGCTGTCACTCAAGCAAGCCCG 419  
288 GGCAGGCTCAGCAATCGTGCAGCTCAATCCCCGGTGGCTGTCACTCAGAGCAAGCTG 347  
420 GTTGTATTGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGGACTACCTAG 479  
348 GCTGTTTTCAGCAATGGGAGCACTATCAGATAAATCAGCAGTGGGAACGAGACTACCTAG 407  
480 GCAATGGTGGTGGTTGACTTGTATGAGGAGGAGCGGAGTTTAACTCGAAAGTAAAC 539  
408 GCAAGCCCTGGTGGTTGACTGCTGATGAGGAGGAGCGGAGTTTAACTCGAGAGCAAGC 467  
540 CTGAAGCTGAAGAGACTGTTTGAACAAGTACACTGGGAACACTTACCAGTGGGTGACA 599  
468 CTGAGCCTGAAGAGACTGTTGTCACAAATACACTGGGAACACTTACAAAGTGGGTGACA 527  
600 CTTATGAGCGCTTAAGAGCTCCATGATCTGGGACTGTACCTGTCATCGGGGCTGGCGAG 659  
528 CTTATGAGCGCTTAAGAGTTCATGATCTGGGACTGTACCTGTCATCGGGGCTGGGAG 587  
660 GGAGAAATAGCTGATCAATCGAAAACGCTGCCATGAAGGGGTGAGTCTTCAAGATTG 719  
588 GCAGGATCAGCTGTACCAATGCAATGCTGCCATGAAGGGGTGAGTCTTCAAGATTG 647  
720 GTGACACCTGGAGGAGACCACTGAGACTGGTGGTGTACATGTTAGAGTGTGTGTCTTG 779  
648 CGACAAAGTGGAGAGGAGCCACATGAGACTGGTGGCTACATGTTAGAGTGTGTGTCTTG 707  
780 GTAAATGAAAAGGAGAAATGGACCTGCAAGCCCATAGCTGAGAAAGTGTGATCATGCTG 839  
708 GAAATGAAAAGGAGAAATGGACCTGCAAGCCCATAGCTGAGAAAGTGTGATCATGCTG 767  
840 CTGGGACTTCTCTATGTGGTCCGAGAAACGTTGGGAGAGCCCTACCAAGCTGGATGATGG 899  
768 CTGGGAGCTCTCTATGCTGGGGGAGACTCTGGGAAAAGCCCTACCAAGCTGGATGATGG 827  
900 TAGATTGTTACTTGGCTGGGAGAGGAGCGGAGCGATCATCTTGCATCTTAGAAATAGAT 959  
828 TGGACTGTACTTGTCTAGGCGAAGGCGAATGGAGCGATCATCTTGCATCTTAGAAATAGAT 887  
960 GCAACGATCAGGACACACCGGAGCATCTCTATAGAAATGGAGACACTCGGACCAAGAGGATA 1019  
888 GCAACGATCAGGACACACCGGAGCATCTCTATAGAAATGGAGACACTCGGACCAAGAGGATA 947  
1020 ATCAGGAGAACTGCTCCAGTGCATCTGCACAGGAAACCGCCGAGGAGAGTGGAGTGTG 1079  
948 ACCGAGGAACTGCTCCAGTGCATCTGCACAGGAAATGGAGAGGAGTGGAGTGTG 1007  
1080 AGAGGACACTCTGTGTCAGACCAATCGAGCGGATCTGGCCCTTCAACGATGTTGCTG 1139  
1008 AGCGACA---TGCTCTACAAAGTGTCTAGCCGGATCTGGCTCTTCACTGATGTCCGAA 1064

1	5324.8	68.3	8315	11	AK090135	AK090135	Mus muscu
2	5187.4	66.5	8329	11	AK090130	AK090130	Mus muscu
3	1342.2	17.2	2067	11	AF130095	AF130095	Homo sapi
4	1043.2	13.4	1201	9	AL542212	AL542212	AL542212

RESULT 1	ACCESSION	REFERENCE
AK090135	VERSION	AUTHORS
LOCUS	KEYWORDS	TITLE
DEFINITION	SOURCE	JOURNAL
	ORGANISM	MEDLINE
		PUBMED
		REFERENCE
		AUTHORS
		TITLE
		JOURNAL
		MEDLINE
		PUBMED

## ALIGNMENTS

AKO90135 8315 bp mRNA linear HTC 20-SEP-2003  
Mus musculus 5 months female bone marrow stroma cell CRU-2028  
SR-4987 cDNA, RIKEN full-length enriched library, clone:G431004D19  
product:fibronectin 1, full insert sequence.  
AKO90135  
AKO90135.1 GI:26105748  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	1	Yamamura, Kuchera, Kocentia, Gerdunoga
AUTHORS		Carninci, P. and Hayashizaki, Y.
TITLE		High-efficiency full-length cDNA cloning
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)
MEDLINE		99279253
PUBMED		10349636
REFERENCE	2	
AUTHORS		Carninci, P., Shibata, Y., Hayatsu, N., Sug
TITLE		Normalization and subtraction of cap-tr
JOURNAL		prepare full-length cDNA libraries for r
MEDLINE		Genome Res. 10 (10), 1617-1630 (2000)
PUBMED		11042159

16	2178.4	27.9	2384	12	US-10-240-425-1289	Sequence 1289, Ap
17	2178.4	27.9	2384	14	US-10-171-311-61	Sequence 61, Appl
18	2178.4	27.9	2384	14	US-10-269-909-23	Sequence 23, Appl
19	2178.4	27.9	2384	16	US-10-447-161-2	Sequence 2, Appl
20	1978.6	25.4	2348	9	US-09-920-300A-1727	Sequence 1727, Ap
21	1978.6	25.4	2348	13	US-10-033-528-1727	Sequence 1727, Ap
22	1978.6	25.4	2348	14	US-10-099-928-1727	Sequence 1727, Ap
23	1895.4	24.3	2127	14	US-10-210-120-49	Sequence 49, Appl
24	1836.4	23.6	4295	15	US-10-144-194A-51	Sequence 51, Appl
25	1561.4	20.0	2841	9	US-09-834-975-1036	Sequence 1026, Ap
26	1400	18.0	2481	10	US-09-775-964-20	Sequence 20, Appl
27	1313.2	16.8	1605	14	US-10-198-846-13899	Sequence 13899, A
28	1202.8	15.4	3069	14	US-10-198-846-13925	Sequence 13925, A
29	1108.2	14.2	1290	14	US-10-198-846-11565	Sequence 11565, A
30	1086.4	13.9	1189	14	US-10-198-846-11556	Sequence 11556, A
31	1020	13.1	1053	9	US-09-934-706-8	Sequence 8, Appl
32	1020	13.1	1224	9	US-09-934-706-16	Sequence 16, Appl
33	1020	13.1	1527	9	US-09-934-706-15	Sequence 15, Appl
34	1016.6	13.0	3069	14	US-10-198-846-13925	Sequence 13925, A
35	990.2	12.7	2841	9	US-09-834-975-1036	Sequence 1026, Ap
36	944.4	12.1	1722	14	US-10-198-846-11562	Sequence 11562, A
37	921.2	11.8	1722	10	US-09-775-964-33	Sequence 33, Appl
38	877.4	11.3	1416	10	US-09-775-964-27	Sequence 27, Appl
39	863.6	11.1	927	14	US-10-198-846-11560	Sequence 11560, A
40	860.8	11.0	1644	10	US-09-775-964-17	Sequence 17, Appl
41	831	10.7	1374	9	US-09-775-964-26	Sequence 26, Appl
42	813.4	10.4	667	9	US-09-925-302-105	Sequence 105, Appl
43	773.8	9.9	777	10	US-09-940-233-3	Sequence 3, Appl
44	734	9.4	963	14	US-10-198-846-11035	Sequence 11035, A
45	703.8	9.0	838	14	US-10-198-846-8931	Sequence 8931, Ap

## RESULT 1

```

US-10-084-817-2
; Sequence 2, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 7795
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 2985802CBL
US-10-084-817-2

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query		Length	DB	ID	Description
		Match	$\tau$				
1	7795	100.0		7795	14	US-10-084-817-2	Sequence 2, Appl
2	7780.6	99.8		8044	14	US-10-240-965-121	Sequence 121, App
3	7774.8	99.7		8137	13	US-10-098-841-8	Sequence 8, Appl
4	7760.2	99.6		8027	16	US-10-447-161-8	Sequence 8, Appl
5	7671.8	98.4		8320	13	US-10-098-841-7	Sequence 7, Appl
6	7614.8	97.7		8062	13	US-10-098-841-5	Sequence 5, Appl
7	7423	95.2		7580	9	US-09-964-824A-574	Sequence 574, App
8	7423	95.2		7580	12	US-10-182-936A-75	Sequence 75, Appl
9	7423	95.2		7580	14	US-10-171-311-63	Sequence 63, Appl
10	7423	95.2		7680	15	US-10-336-031B-69	Sequence 69, Appl
11	7423	95.2		7680	15	US-10-374-979-75	Sequence 75, Appl
12	7418.2	95.2		7705	16	US-10-447-161-4	Sequence 4, Appl
13	7224.8	92.7		7867	13	US-10-098-841-6	Sequence 6, Appl
14	5320.6	68.3		8329	9	US-09-917-800A-1731	Sequence 1731, Ap
15	5320.6	68.3		8329	15	US-10-191-803-113	Sequence 113, App

121 y CCGCTTCCCAACCTCTGGCCCCCACCCTTCTTGGAGGACAAACCCCGGGAGGCATTAG 180  
121 b CCGCTTCCCAACCTCTGGCCCCCACCCTTCTTGGAGGACAAACCCCGGGAGGCATTAG 180  
181 y AAGGATTTTCCCGAGGTTGGGAAGGAAACAACTTGGTGGCAACTTGCCTCCCGGT 240  
181 b AAGGATTTTCCCGAGGTTGGGAAGGAAACAACTTGGTGGCAACTTGCCTCCCGGT 240  
241 y GGGGGCTCTCTCCCCCACCCTCTCAACATGCTTAGGGGTCCGGGGCCCGGCTCTGCT 300  
241 b GGGGGCTCTCTCCCCCACCCTCTCAACATGCTTAGGGGTCCGGGGCCCGGCTCTGCT 300  
301 y GCTGGCCCTCTGTGCTGGGACACAGCGTGGCTTCCACGGAGGCTTGAAGAGCAAGAG 360  
301 b GCTGGCCCTCTGTGCTGGGACACAGCGTGGCTTCCACGGAGGCTTGAAGAGCAAGAG 360  
361 y GAGGCTCAGCAAAATGGTTGAGCCCCAGTCCCGGTGGCTGTCACTCAAAAGCAAGCCGG 420  
361 b GAGGCTCAGCAAAATGGTTGAGCCCCAGTCCCGGTGGCTGTCACTCAAAAGCAAGCCGG 420  
421 y TTGTTATGACATGGAAGAACATATCAGATAAATCAACGTGGAGCGGACCTACTAGG 480  
421 b TTGTTATGACATGGAAGAACATATCAGATAAATCAACGTGGAGCGGACCTACTAGG 480  
481 y CAATGGCTTGGTTGTACTTGTATGAGAGAGCGAGGTTTAACTGCGAAAGTAAACC 540  
481 b CAATGGCTTGGTTGTACTTGTATGAGAGAGCGAGGTTTAACTGCGAAAGTAAACC 540  
541 y TGAAGCTGAAGAGACTTGTCTTGAAGAATGACCTGGGAACATTAACGAGTGGGTGAC 600  
541 b TGAAGCTGAAGAGACTTGTCTTGAAGAATGACCTGGGAACATTAACGAGTGGGTGAC 600  
601 y TTATGAGCGTCTTAAAGACTCATGATCTGGGACTGTACTGCATCGGGCTGGCCGAGG 660  
601 b TTATGAGCGTCTTAAAGACTCATGATCTGGGACTGTACTGCATCGGGCTGGCCGAGG 660  
661 y GAGAAATGCTGTACATCGCAAAACCGTCCATGAAGGGGTCTAGTCTCTACAAGATTGG 720  
661 b GAGAAATGCTGTACATCGCAAAACCGTCCATGAAGGGGTCTAGTCTCTACAAGATTGG 720  
721 y TGACACTGGAGGAGACCATGAGACTGGTGGTTACATGTTAGAGTGTGTGTTGG 780  
721 b TGACACTGGAGGAGACCATGAGACTGGTGGTTACATGTTAGAGTGTGTGTTGG 780  
781 y TAATGGAAGAGAGATGCAAGCTGCAAGCCCATAGCTGAGAAAGTGTGTTGATCATGCTGC 840  
781 b TAATGGAAGAGAGATGCAAGCTGCAAGCCCATAGCTGAGAAAGTGTGTTGATCATGCTGC 840  
841 y TGGGACTTCTATGTGGTGGGAAACGTTGGAGAGCCCTACCAAGCTGGATGATGTT 900  
841 b TGGGACTTCTATGTGGTGGGAAACGTTGGAGAGCCCTACCAAGCTGGATGATGTT 900  
901 y AGATTGCTACTTGGCTGGGAGAGGACGCGGACGATCACTTGCATCTTCTAGAAATAGATG 960  
901 b AGATTGCTACTTGGCTGGGAGAGGACGCGGACGATCACTTGCATCTTCTAGAAATAGATG 960  
961 y CAACGATCAGGACACAAAGACATCTTATAGATTTGGAGACATCTGGAGCAAGAGGATAA 1020  
961 b CAACGATCAGGACACAAAGACATCTTATAGATTTGGAGACATCTGGAGCAAGAGGATAA 1020  
1021 y TCAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGCGGAGAGAGTGGAAAGTGA 1080  
1021 b TCAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGCGGAGAGAGTGGAAAGTGA 1080  
1081 y GAGGACACCTCTGTGTCAGACACATCGAGCGGATCTGGCCCTCTCAACGATGTTCTGTC 1140  
1081 b GAGGACACCTCTGTGTCAGACACATCGAGCGGATCTGGCCCTCTCAACGATGTTCTGTC 1140  
1141 y AGCTGTTTACCAACCGAGCTTCAACCCAGGCTCTCTCCCTATGGCACTGTGTACAGA 1200  
1141 b AGCTGTTTACCAACCGAGCTTCAACCCAGGCTCTCTCCCTATGGCACTGTGTACAGA 1200  
1201 y CAGTGTGTGGTCTACTCTGTGGGGATGTCAGTGGTGAAGACAAAGGAATAAAGCAAAAT 1260

1201 Db CAGTGTGTGGTCTACTCTGTGGGATGCACTGGCTGAAGACACAAGGAATAAAGCAAAAT 1260  
1261 QY GCCTTGCAGCTGCTGGCAACGAGCTCAGCTGCAAGAGACAGCTGTAAACCCAGACTTA 1320  
1261 Db GCCTTGCAGCTGCTGGCAACGAGCTCAGCTGCAAGAGACAGCTGTAAACCCAGACTTA 1320  
1321 QY CGGTGGCAACTCAAAATGGAGAGCCATGTGTCTTACCAATTCACCTAATGGCAGGACGTT 1380  
1321 Db CGGTGGCAACTCAAAATGGAGAGCCATGTGTCTTACCAATTCACCTAATGGCAGGACGTT 1380  
1381 QY CTACTCTGACACACAGAGAGGCGAGAGCGGACATCTTTGGTGCAGCAAACTTCGAA 1440  
1381 Db CTACTCTGACACACAGAGAGGCGAGAGCGGACATCTTTGGTGCAGCAAACTTCGAA 1440  
1441 QY TTATGAGCAGGACAGAAATPACTCTTTTGCACAGACCACTGTTTGGTTCAGACTCG 1500  
1441 Db TTATGAGCAGGACAGAAATPACTCTTTTGCACAGACCACTGTTTGGTTCAGACTCG 1500  
1501 QY AGAGGAAATTCCTTGGTGGCTTGGCCCTTGGCCCTTCCCTTATACAAACCACTTA 1560  
1501 Db AGAGGAAATTCCTTGGTGGCTTGGCCCTTGGCCCTTCCCTTATACAAACCACTTA 1560  
1561 QY CACTGATTGCACTTCTGAGGGCAGAAAGAGACAACTGAAGTGGTGGGACCAACAGAA 1620  
1561 Db CACTGATTGCACTTCTGAGGGCAGAAAGAGACAACTGAAGTGGTGGGACCAACAGAA 1620  
1621 QY CTATGATCCGACACAGAAATTTGGTTCGCCCTTGGCCCTTCCCTTATACAAACCACTTA 1680  
1621 Db CTATGATCCGACACAGAAATTTGGTTCGCCCTTGGCCCTTCCCTTATACAAACCACTTA 1680  
1681 QY AACCAATGAAGGGGTCTATGTACCGCATTTGGAGATCAGTGGGATGAAGTGGTGGGACCACTTA 1740  
1681 Db AACCAATGAAGGGGTCTATGTACCGCATTTGGAGATCAGTGGGATGAAGTGGTGGGACCACTTA 1740  
1741 QY TCACATGATGAGTGCAGCTGTGTGGGAAATGGTTCGCGGAAATGGGATGACATGCTTTCCTTA 1800  
1741 Db TCACATGATGAGTGCAGCTGTGTGGGAAATGGTTCGCGGAAATGGGATGACATGCTTTCCTTA 1800  
1801 QY CTGCGAGCTTCAGAGATCAGTGCATTTGATGACATCACTTACAATGTGAACGACACAT 1860  
1801 Db CTGCGAGCTTCAGAGATCAGTGCATTTGATGACATCACTTACAATGTGAACGACACAT 1860  
1861 QY CCAAGCGTTCATGAAGGGGCACTGCTGAACCTGATCATGCTTCGGTTCAGGCTCGGG 1920  
1861 Db CCAAGCGTTCATGAAGGGGCACTGCTGAACCTGATCATGCTTCGGTTCAGGCTCGGG 1920  
1921 QY CAGTGAAGTGTGATCCCGTGCACCAATGCCAGGATTCAGAGACTGGGACGTTTTATCA 1980  
1921 Db CAGTGAAGTGTGATCCCGTGCACCAATGCCAGGATTCAGAGACTGGGACGTTTTATCA 1980  
1981 QY AATTGGAGATTCATGGGAGAGTATGTGCATGTGTGATACCATACCATGCTACTGCTATGG 2040  
1981 Db AATTGGAGATTCATGGGAGAGTATGTGCATGTGTGATACCATACCATGCTACTGCTATGG 2040  
2041 QY CCGTGGCATTTGGGAGTGGCACTTGCACCACTTTACAGACTTATCCAAAGCTCAAGTGGTCC 2100  
2041 Db CCGTGGCATTTGGGAGTGGCACTTGCACCACTTTACAGACTTATCCAAAGCTCAAGTGGTCC 2100  
2101 QY TGTGAAAGTATTTTATCACTGAGACTCCGAGTCAAGCCAACTCCCACTCCCATCCAGTGGAA 2160  
2101 Db TGTGAAAGTATTTTATCACTGAGACTCCGAGTCAAGCCAACTCCCACTCCCATCCAGTGGAA 2160  
2161 QY TGACACACAGGATCTCACTATTTCCAGTACATTTCCAGTGGGAGACCTTAAATTCCT 2220  
2161 Db TGACACACAGGATCTCACTATTTCCAGTACATTTCCAGTGGGAGACCTTAAATTCCT 2220  
2221 QY AGGCGGTTGAAGAGAGTACCATACAGGCCACTTAACTCCTACCATCAAGGCGCT 2280  
2221 Db AGGCGGTTGAAGAGAGTACCATACAGGCCACTTAACTCCTACCATCAAGGCGCT 2280  
2281 QY GAAGCTGTGTGGTATATCAAGGGCAGCTCATCAGATCCAGCAGTACGGCCACCAAGA 2340

Db 2281 GAAGCCTGGTGTGTATACGAGGGCCAGCTCATCAGCATCAGCAGTACGGCCACCAAGA 2340  
2y AGTGACTCGCTTTGACTTTACCAACACACAGCACACCACTGTGACCAACACCGT 2400  
Db 2341 AGTGACTCGCTTTGACTTTACCAACACACAGCACACCACTGTGACCAACACCGT 2400  
2y GACAGGAGAGAGACTCCCTTTTCTCTCTGTGGCCACTTCTGAATCTGTGACCGAAT 2460  
Db 2401 GACAGGAGAGAGACTCCCTTTTCTCTCTGTGGCCACTTCTGAATCTGTGACCGAAT 2460  
2y CACAGCCAGTAGCTTTGTGGTCTCTCTGGGTCTCAGCTTCCGACACCGTGTGCGGATTCG 2520  
Db 2461 CACAGCCAGTAGCTTTGTGGTCTCTCTGGGTCTCAGCTTCCGACACCGTGTGCGGATTCG 2520  
2y GGTGGATATAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGATCTTCCAAAGCAC 2580  
Db 2521 GGTGGATATAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGATCTTCCAAAGCAC 2580  
2y AGCCACTTCTGTGAACATCCCTGACCTGCTTCTGGCCGAAATACATTTGTAATGTCTA 2640  
Db 2581 AGCCACTTCTGTGAACATCCCTGACCTGCTTCTGGCCGAAATACATTTGTAATGTCTA 2640  
2y TCAGATATCTGAGGATGGGAGAGAGTTGATCTGTCTACTTTCACAAACAGCGCC 2700  
Db 2641 TCAGATATCTGAGGATGGGAGAGAGTTGATCTGTCTACTTTCACAAACAGCGCC 2700  
2y TGATGCCCTCTCTGACCCGACTGGACCAAGTTGATGACACCTCAATTTGTTTCGCTG 2760  
Db 2701 TGATGCCCTCTCTGACCCGACTGGACCAAGTTGATGACACCTCAATTTGTTTCGCTG 2760  
2y GAGCAGACCCGAGGCTCCATTCACAGGGTACAGAAATGATTCGGCCATCAGTAGAAG 2820  
Db 2761 GAGCAGACCCGAGGCTCCATTCACAGGGTACAGAAATGATTCGGCCATCAGTAGAAG 2820  
2y TAGCAGCAGACAACTCAACCTTCTGAACTGCAAACTCCGTGACCTCAGTGACTTGCA 2880  
Db 2821 TAGCAGCAGACAACTCAACCTTCTGAACTGCAAACTCCGTGACCTCAGTGACTTGCA 2880  
2y ACCTGGTTCAGTATACATCACTATCTATCTGTGGAAGAAATCAAGAAAGTACAC 2940  
Db 2881 ACCTGGTTCAGTATACATCACTATCTATCTGTGGAAGAAATCAAGAAAGTACAC 2940  
2y TGTTGTCAATCAAGAAACCACTGGCACCCACGCTCAGATACAGTCCCTCCAG 3000  
Db 2941 TGTTGTCAATCAAGAAACCACTGGCACCCACGCTCAGATACAGTCCCTCCAG 3000  
2y GGACCTGCAGTTTGTGGAAGTGACAGAGTGAAGGTCAACATCATGTGGACACCGCCTGA 3060  
Db 3001 GGACCTGCAGTTTGTGGAAGTGACAGAGTGAAGGTCAACATCATGTGGACACCGCCTGA 3060  
2y GAGTGCAGTACCGGTACCGTGTGGATGTGATCCCGTCAACCTGGCGAGCACGG 3120  
Db 3061 GAGTGCAGTACCGGTACCGTGTGGATGTGATCCCGTCAACCTGGCGAGCACGG 3120  
2y GCAGAGGCTGCCATCAGCAGGAACACCTTTGAGAAAGTCAACCGGCTGCCCCGGGT 3180  
Db 3121 GCAGAGGCTGCCATCAGCAGGAACACCTTTGAGAAAGTCAACCGGCTGCCCCGGGT 3180  
2y CACCTATTACTTCAAGTCTTTGAGTGAGCCATGGGAGGAGAGAGAGCTCTGACTGC 3240  
Db 3181 CACCTATTACTTCAAGTCTTTGAGTGAGCCATGGGAGGAGAGAGAGCTCTGACTGC 3240  
2y TCAACAGACAAACCAACTGGATGCTCCCACTAACTCCAGTTTCTCAATGAACCTGATTC 3300  
Db 3241 TCAACAGACAAACCAACTGGATGCTCCCACTAACTCCAGTTTCTCAATGAACCTGATTC 3300  
2y TACTGTCTGTGTAGATGGAATCTCACTCTGGGGCCAGATTAACAGGATACCGATGACCGT 3360  
Db 3301 TACTGTCTGTGTAGATGGAATCTCACTCTGGGGCCAGATTAACAGGATACCGATGACCGT 3360  
2y GGGCTTACCCGAGAGAGACAGCCAGCAGTACAAATGTGGGTCCCTGTCTCCAAAGTA 3420  
Db 3361 GGGCTTACCCGAGAGAGACAGCCAGCAGTACAAATGTGGGTCCCTGTCTCCAAAGTA 3420

Qy 3421 CCACCTGAGGAATCTGACGCTGCACTCTGAGTACACCGTATCCCTCGTGGCCATAAAGGG 3480  
Db 3421 CCACCTGAGGAATCTGACGCTGCACTCTGAGTACACCGTATCCCTCGTGGCCATAAAGGG 3480  
Qy 3481 CAACCAAGAGAGCCCAAGCCACTGGAGTCTTTACACACACTGACAGCCTGGAGCTCTAT 3540  
Db 3481 CAACCAAGAGAGCCCAAGCCACTGGAGTCTTTACACACACTGACAGCCTGGAGCTCTAT 3540  
Qy 3541 TCACCTTACAAACAACGAGGTGACTGAGACACACCATTTGTGATCATGAGCGCTGCTCC 3600  
Db 3541 TCACCTTACAAACAACGAGGTGACTGAGACACACCATTTGTGATCATGAGCGCTGCTCC 3600  
Qy 3601 AGAATTTGTTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGACACACGAGAAGT 3660  
Db 3601 AGAATTTGTTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGACACACGAGAAGT 3660  
Qy 3661 GACTTCAGACTCAGGAAGCATCGTTGTGTCCGGCTTGACTCCAGGAGTAGAATACGTCTA 3720  
Db 3661 GACTTCAGACTCAGGAAGCATCGTTGTGTCCGGCTTGACTCCAGGAGTAGAATACGTCTA 3720  
Qy 3721 CACCATCCAAGTCTGAGAGATGGACAGAAAGAGATGCCAAATTTGTAACAAAGTGGT 3780  
Db 3721 CACCATCCAAGTCTGAGAGATGGACAGAAAGAGATGCCAAATTTGTAACAAAGTGGT 3780  
Qy 3781 GACACCATTTGTCTCCACCAACAACTTGATCTGGAGGCAAACTCCGTGACCTGGAGTGT 3840  
Db 3781 GACACCATTTGTCTCCACCAACAACTTGATCTGGAGGCAAACTCCGTGACCTGGAGTGT 3840  
Qy 3841 CACAGTCTCTGGGAGAGAGACCAACCCAGACATTAAGTGGTTATAGAAATACCAACAC 3900  
Db 3841 CACAGTCTCTGGGAGAGAGACCAACCCAGACATTAAGTGGTTATAGAAATACCAACAC 3900  
Qy 3901 CCTACAAACGGCCAGCAGGGAATTTCTTTGGAAGAGTGTCCAATGCTGATCAGAGTCT 3960  
Db 3901 CCTACAAACGGCCAGCAGGGAATTTCTTTGGAAGAGTGTCCAATGCTGATCAGAGTCT 3960  
Qy 3961 CTGCACTTTTGTATACCTGAGTCCCGGCTGGAGTACAAATGTCAGTGTTTACACTGTCAA 4020  
Db 3961 CTGCACTTTTGTATACCTGAGTCCCGGCTGGAGTACAAATGTCAGTGTTTACACTGTCAA 4020  
Qy 4021 GGATGACAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGTTCCTCCCTCCAC 4080  
Db 4021 GGATGACAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGTTCCTCCCTCCAC 4080  
Qy 4081 TGACCTTGCATTCACCAACATTTGGTCCAGACACCATGCGTGTCACTGGGTCCACCCCC 4140  
Db 4081 TGACCTTGCATTCACCAACATTTGGTCCAGACACCATGCGTGTCACTGGGTCCACCCCC 4140  
Qy 4141 ATCCATTTGATTTAACCACTTCTGGTGGTTACTCACTGTGAAAATAGGAAGATGT 4200  
Db 4141 ATCCATTTGATTTAACCACTTCTGGTGGTTACTCACTGTGAAAATAGGAAGATGT 4200  
Qy 4201 TGCAGAGTTGTCAAATTTCTCTTCAGACAAATGCACTGGTCTTTAAACAAATCTCTCCCTGG 4260  
Db 4201 TGCAGAGTTGTCAAATTTCTCTTCAGACAAATGCACTGGTCTTTAAACAAATCTCTCCCTGG 4260  
Qy 4261 TACAGATATGTAGTGTCTCCAGTGTCTTACGAAACAAATGAGAGACACCTCTTAG 4320  
Db 4261 TACAGATATGTAGTGTCTCCAGTGTCTTACGAAACAAATGAGAGACACCTCTTAG 4320  
Qy 4321 AGGAAGACAGAAAACAGGTCTTTGATTTCCCAACTGGCAATGACATTTCTGATATTACTGC 4380  
Db 4321 AGGAAGACAGAAAACAGGTCTTTGATTTCCCAACTGGCAATGACATTTCTGATATTACTGC 4380  
Qy 4381 GAACTCTTTTACTGTGCACTGGATTGCTCTCGAGCCACCATCACTGGCTACAGATCCG 4440  
Db 4381 GAACTCTTTTACTGTGCACTGGATTGCTCTCGAGCCACCATCACTGGCTACAGATCCG 4440  
Qy 4441 CCATCATCCCGAGCACTTCAGTGGGAGACCTCGAGAAAGATCGGGTGCCTCTCTCGAA 4500  
Db 4441 CCATCATCCCGAGCACTTCAGTGGGAGACCTCGAGAAAGATCGGGTGCCTCTCTCGAA 4500

4501	TTCCATCACCCCTACCAACCTCACTCCAGCAGAGTAGTGTGATCAGCATCGTTGCTCT	4560
4501	TTCCATCACCCCTACCAACCTCACTCCAGCAGAGTAGTGTGATCAGCATCGTTGCTCT	4560
4561	TATGGCAGAGAGGAAGTCCCTTATTTGATTGGCCCAACCAATCAACAGTTTCTTGATGTTCC	4620
4561	TATGGCAGAGAGGAAGTCCCTTATTTGATTGGCCCAACCAATCAACAGTTTCTTGATGTTCC	4620
4621	GAGGGACCTGGAAAGTTGTTGCTGCGACCCCCACCCAGCTTCACTCAGCTGGGATGCTCC	4680
4621	GAGGGACCTGGAAAGTTGTTGCTGCGACCCCCACCCAGCTTCACTCAGCTGGGATGCTCC	4680
4681	TGCTGTCAAGTCAGATATTTACAGGATCACTTACGGAGAAACAGGAGGAAATGACCCCTCT	4740
4681	TGCTGTCAAGTCAGATATTTACAGGATCACTTACGGAGAAACAGGAGGAAATGACCCCTCT	4740
4741	CCAGGAGTTCACTGTGCTGGGAGCAAGTCTACAGTACCATCAGCGGGCTTAAACCTGCG	4800
4741	CCAGGAGTTCACTGTGCTGGGAGCAAGTCTACAGTACCATCAGCGGGCTTAAACCTGCG	4800
4801	AGTTGATTATACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAGCCCCGCAAGCAG	4860
4801	AGTTGATTATACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAGCCCCGCAAGCAG	4860
4861	CAAGCCAAATTTCCAATTAATTCGGAAACAGAAATTTGACAAACCATCCACAGTCAAACTGAC	4920
4861	CAAGCCAAATTTCCAATTAATTCGGAAACAGAAATTTGACAAACCATCCACAGTCAAACTGAC	4920
4921	CGATGTTCAGGACAAACAGCATATTAAGTGTCAAGTGGCTTCAAGTTCCCTGTTACTGG	4980
4921	CGATGTTCAGGACAAACAGCATATTAAGTGTCAAGTGGCTTCAAGTTCCCTGTTACTGG	4980
4981	TTACAGAGTAAACACCACTCCCAAAATTTGACAGGACCAACAAAACATAAACTGCAAG	5040
4981	TTACAGAGTAAACACCACTCCCAAAATTTGACAGGACCAACAAAACATAAACTGCAAG	5040
5041	TCAGATCAAAACAGAAATGACTATTGAAGGCTTTCAGGCCACACAGTGGAGTATGTGTTAG	5100
5041	TCAGATCAAAACAGAAATGACTATTGAAGGCTTTCAGGCCACACAGTGGAGTATGTGTTAG	5100
5101	TGTCTATGCTCAGAAATCCCAAGCGGAGAGTCAAGCCTCTGTTTTCAGACTGCAGTAAACCA	5160
5101	TGTCTATGCTCAGAAATCCCAAGCGGAGAGTCAAGCCTCTGTTTTCAGACTGCAGTAAACCA	5160
5161	CATTGATCGCCCTAAAGGACTGGCATTCACTGATGTGGATGTGATTTCCATCAAAATTCG	5220
5161	CATTGATCGCCCTAAAGGACTGGCATTCACTGATGTGGATGTGATTTCCATCAAAATTCG	5220
5221	TTGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGTCAGCTTCTCAGAGCCCTGAGGA	5280
5221	TTGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGTCAGCTTCTCAGAGCCCTGAGGA	5280
5281	TGGAATCCATGAGCTATTCCCTGCACCTGATGTGTGAAGAAGACACTGCAGAGCTGCAAGG	5340
5281	TGGAATCCATGAGCTATTCCCTGCACCTGATGTGTGAAGAAGACACTGCAGAGCTGCAAGG	5340
5341	CCTCAGACCGGGTTCTGAGTACACAGTCAGTGTGGTTGCTTGCACGATGATGGAGAG	5400
5341	CCTCAGACCGGGTTCTGAGTACACAGTCAGTGTGGTTGCTTGCACGATGATGGAGAG	5400
5401	CCAGCCCTGATTGGAAACCCAGTCCACAGCTATTTCCTGCAACCACTGACCTGAAGTTTCA	5460
5401	CCAGCCCTGATTGGAAACCCAGTCCACAGCTATTTCCTGCAACCACTGACCTGAAGTTTCA	5460
5461	TCAGGTCAACCCCAAGACCTTGAGCGGCCAGTGACACCAACCAATGTTTCACTCTCTG	5520
5461	TCAGGTCAACCCCAAGACCTTGAGCGGCCAGTGACACCAACCAATGTTTCACTCTCTG	5520
5521	ATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAATCAACCTTGC	5580
5521	ATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAATCAACCTTGC	5580
5581	TCCTGACAGCTCATCCGTGGTTGTATCAGGACTTATGTTGGTGCACCAATATGAAGTAG	5640

5581	TCTTGACAGCTCATCCGTGGTGTGTATCAGGACTTATGGTGGCCACCAATATGAAGTGAG	5640
5641	TGCTCTATGCTCTTTAAGGACACTTTTGACAAAGCAGACCAGCTCAGGGAGTTGTCAACCACTCT	5700
5641	TGCTCTATGCTCTTTAAGGACACTTTTGACAAAGCAGACCAGCTCAGGGAGTTGTCAACCACTCT	5700
5701	GGAGAAATGTGAGCCCAACAGAGGGCTCGTGTGACAGATGCTACTGACGACCACCATCAC	5760
5701	GGAGAAATGTGAGCCCAACAGAGGGCTCGTGTGACAGATGCTACTGACGACCACCATCAC	5760
5761	CATTAGCTTGAAGAACCAAGAGCTCGAGACGATCACTGGCCTTCAAGTTGATGCGTGTCCAGC	5820
5761	CATTAGCTTGAAGAACCAAGAGCTCGAGACGATCACTGGCCTTCAAGTTGATGCGTGTCCAGC	5820
5821	CAATGGCCAGACTCCAAATCAGAGAACCATCAAGCCAGATGTTCAGAGCTACACCATCAC	5880
5821	CAATGGCCAGACTCCAAATCAGAGAACCATCAAGCCAGATGTTCAGAGCTACACCATCAC	5880
5881	TGGCTTTACAAACGAGCACTCACTACAAAGATCTACCTGTACACCTTTGAATGACAAATGCTCG	5940
5881	TGGCTTTACAAACGAGCACTCACTACAAAGATCTACCTGTACACCTTTGAATGACAAATGCTCG	5940
5941	GAGCTCCCTGTGTCTATGAGAGCTCCCACTGTCATGTATGCAATCATCAATCCAACTGCGTTT	6000
5941	GAGCTCCCTGTGTGTCTATGAGAGCTCCCACTGTCATGTATGCAATCATCAATCCAACTGCGTTT	6000
6001	CTTGCCACCAACCAACCAATTCCTTGTGTGTATCATGGACGCGCCACGCTGCCAGATTTAC	6060
6001	CTTGCCACCAACCAACCAATTCCTTGTGTGTATCATGGACGCGCCACGCTGCCAGATTTAC	6060
6061	CGGCTACATCATCAAGTATGAGAGGCTGGGTCTCTCCAGAGAAGTGTGCTTCCCTCGGCC	6120
6061	CGGCTACATCATCAAGTATGAGAGGCTGGGTCTCTCCAGAGAAGTGTGCTTCCCTCGGCC	6120
6121	CGGCTTGTGTCTACAGAGGCTACTATTACTGGCTTGGAAACGGGAAACCGAATATACAAT	6180
6121	CGGCTTGTGTCTACAGAGGCTACTATTACTGGCTTGGAAACGGGAAACCGAATATACAAT	6180
6181	TTATGTCTATTGGCTTGAAGATTAATCAGAGAGCGAGCCCTGTATGGAAGGAAAGAGAC	6240
6181	TTATGTCTATTGGCTTGAAGATTAATCAGAGAGCGAGCCCTGTATGGAAGGAAAGAGAC	6240
6241	AGACGAGCTTCCCAGCTGCTGCTTCAACACCCCAATCTTCAATGAGACGAGATCTT	6300
6241	AGACGAGCTTCCCAGCTGCTGCTTCAACACCCCAATCTTCAATGAGACGAGATCTT	6300
6301	GGATGTTCTCTTCCACAGTTTCAAAGAGCCCTTTGCTCACCCACCTCTGGGTATGACACTGG	6360
6301	GGATGTTCTCTTCCACAGTTTCAAAGAGCCCTTTGCTCACCCACCTCTGGGTATGACACTGG	6360
6361	AAATGATATCAGTTCCTGGCATCTTGTGTAGCAAAACCGATGTTTGGGCAACAAATGAT	6420
6361	AAATGATATCAGTTCCTGGCATCTTGTGTAGCAAAACCGATGTTTGGGCAACAAATGAT	6420
6421	CTTTGAGGAACATGTTTATAGCGGACCAACCGCCCAACAAACCGCCCAATTAAGGCA	6480
6421	CTTTGAGGAACATGTTTATAGCGGACCAACCGCCCAACAAACCGCCCAATTAAGGCA	6480
6481	TAGGCCAAGACCATACCOCGGAATGTAGGACAAAGAGTCTCTCTCTCAGACCAACATCTC	6540
6481	TAGGCCAAGACCATACCOCGGAATGTAGGACAAAGAGTCTCTCTCTCAGACCAACATCTC	6540
6541	ATGGGCCCATTTCCAGGACCTTCTGAGTACATCAATTCATGTCATCTCTTGGGCACTGA	6600
6541	ATGGGCCCATTTCCAGGACCTTCTGAGTACATCAATTCATGTCATCTCTTGGGCACTGA	6600
6601	TGAAGAACCCCTTACAGTTTCAAGGTTCTCGAACTTCTACAGTGGCACTCTGACAGGCTT	6660
6601	TGAAGAACCCCTTACAGTTTCAAGGTTCTCGAACTTCTACAGTGGCACTCTGACAGGCTT	6660
6661	CACGAGGTGCCACTTACACATCATAGTGGAGGCACTTGAAAGCCAGCAGAGGCATAA	6720



6661 CACCAGAGGTGCCACTACAAATCATATAGTGGAGGCACTGAAGAACACAGCAGAGGCATAA 6720  
6721 GGTTCGGGAGAGGCTTGTACCGTGGGCACTCTGTCAACGAGAGGCTTGAACCAACCTAC 6780  
6721 GGTTCGGGAGAGGCTTGTACCGTGGGCACTCTGTCAACGAGAGGCTTGAACCAACCTAC 6780  
6781 GGATGACTCGTCTTGAACCCCTACACAGTGTCCCATTTATGCCGTTGGAGATGAGTGGGA 6840  
6781 GGATGACTCGTCTTGAACCCCTACACAGTGTCCCATTTATGCCGTTGGAGATGAGTGGGA 6840  
6841 ACGAATGCTGAATCAGGCTTAACTGTGTGCGAGTGTCTTGGAGTGTGCA 6900  
6841 ACGAATGCTGAATCAGGCTTAACTGTGTGCGAGTGTCTTGGAGTGTGCA 6900  
6901 TTTTCAGATGTCATTCATCTAGATGGTGCATGACAAATGGTGTGAATTTGGA 6960  
6901 TTTTCAGATGTCATTCATCTAGATGGTGCATGACAAATGGTGTGAATTTGGA 6960  
6961 GAATGGGACCGTCAGGAGAAATGGCCAGATGATGATGTCACATGTCTTGGGACGG 7020  
6961 GAATGGGACCGTCAGGAGAAATGGCCAGATGATGATGTCACATGTCTTGGGACGG 7020  
7021 AAAAGGAGAAATTCAGTGTGACCTCATGAGGCAACGTTGTACGATGATGGGAGACATA 7080  
7021 AAAAGGAGAAATTCAGTGTGACCTCATGAGGCAACGTTGTACGATGATGGGAGACATA 7080  
7081 CCAGTAGGAGAAACAGTGGGAGAGAAATATCTGGTGCCATTTGCTCTGCGACATGCTT 7140  
7081 CCAGTAGGAGAAACAGTGGGAGAGAAATATCTGGTGCCATTTGCTCTGCGACATGCTT 7140  
7141 TGGAGGCCAGCGGGCTGGCGCTGTGACAACTGCGCAGACCTGGGGGTGAACCCAGTCC 7200  
7141 TGGAGGCCAGCGGGCTGGCGCTGTGACAACTGCGCAGACCTGGGGGTGAACCCAGTCC 7200  
7201 CGAAGGCACTACTGCGCAGTCTTCAACACAGTATTTCTCAGAGATACCATCAGAGAACAA 7260  
7201 CGAAGGCACTACTGCGCAGTCTTCAACACAGTATTTCTCAGAGATACCATCAGAGAACAA 7260  
7261 CACTAAATGTTAATTCGCAATTTAGTGTCTCATGCTTTAGATGTACAGGCTGACAGAGA 7320  
7261 CACTAAATGTTAATTCGCAATTTAGTGTCTCATGCTTTAGATGTACAGGCTGACAGAGA 7320  
7321 AGATTCGCGAGAGTAAATCATCTTTTCCAAATCAGAGGAAACAGATGTCTCTGCGCAAG 7380  
7321 AGATTCGCGAGAGTAAATCATCTTTTCCAAATCAGAGGAAACAGATGTCTCTGCGCAAG 7380  
7381 ATCCATCTAAATCTGAGTGTAGTGTAGCAGCCAGCTTAGAGTCTTCTTTCTTTTAA 7440  
7381 ATCCATCTAAATCTGAGTGTAGTGTAGCAGCCAGCTTAGAGTCTTCTTTCTTTTAA 7440  
7441 ATCCATCTAAATCTGAGTGTAGTGTAGCAGCCAGCTTAGAGTCTTCTTTCTTTTAA 7500  
7441 ATCCATCTAAATCTGAGTGTAGTGTAGCAGCCAGCTTAGAGTCTTCTTTCTTTTAA 7500  
7501 ACCCTGGAGTTTCTGAGGTTTCTATAAATGAGGCTGCATATGCTGCTGCTGCT 7560  
7501 ACCCTGGAGTTTCTGAGGTTTCTATAAATGAGGCTGCATATGCTGCTGCTGCTGCT 7560  
7561 TCGAAGTATTCATACCGCTCAGTATTTTAAATGAAGTGAATCTAAGATTTGGTTGGGA 7620  
7561 TCGAAGTATTCATACCGCTCAGTATTTTAAATGAAGTGAATCTAAGATTTGGTTGGGA 7620  
7621 TCAATAGGAAACGATATGAGCCCAACCAAGATGCAAAATGTTTGAATGATATGACCAA 7680  
7621 TCAATAGGAAACGATATGAGCCCAACCAAGATGCAAAATGTTTGAATGATATGACCAA 7680  
7681 ATTTTAAAGTAGGAAAGTACCCCAACACCTTCGCTTTTCACTTAACTGCTGCGCCGCAAT 7740  
7681 ATTTTAAAGTAGGAAAGTACCCCAACACCTTCGCTTTTCACTTAACTGCTGCGCCGCAAT 7740  
7741 ACTGAGGAAACAGCATGATCTTTTACTGTGATATTTTAAATATCCACAGTACT 7795  
7741 ACTGAGGAAACAGCATGATCTTTTACTGTGATATTTTAAATATCCACAGTACT 7795

RESULT 2

US-10-240-965-121  
; Sequence 121, Application US/10240965  
; Publication No. US20030165924A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: SHIFFMAN, Dov  
; APPLICANT: SOMOGYI, Roland  
; APPLICANT: LAWN, Richard M.  
; APPLICANT: SEILHAMER, Jeffrey J.  
; APPLICANT: PORTER, Gordon J.  
; APPLICANT: MIKITA, Thomas  
; APPLICANT: TAL, Julie  
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION  
; FILE REFERENCE: PA-0025 PCT  
; CURRENT APPLICATION NUMBER: US/10/240,965  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: 60/195,106  
; PRIOR FILING DATE: 2000-04-05  
; NUMBER OF SEQ ID NOS: 276  
; SOFTWARE: PERL Program  
; SEQ ID NO 121  
; LENGTH: 8044  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030165924A1 427813.14  
US-10-240-965-121

Query Match 99.8%; Score 7780.6; DB 14; Length 8044;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 7786; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 1 CGGCTCGAGCCGGCTGTGTGTCACAGGGGGAGAGAGGAAACCCAGGCGGAGCGGAA 60  
DB 1 CGGCTCGAGCCGGCTGTGTGTCACAGGGGGAGAGAGGAAACCCAGGCGGAGCGGAA 60  
QY 61 GAGGGACCTGACGAGCACAACATTTCTGTGTCTCTGTGATCCCTTCTGTCCCTCCACCCGT 120  
DB 61 GAGGGACCTGACGAGCACAACATTTCTGTGTCTCTGTGATCCCTTCTGTCCCTCCACCCGT 120  
QY 121 CCCCTTCCCACCCCTCTGCCGCCACCTTTCTTGGAGGCGACAACCCCGGGAGGCGATTAG 180  
DB 121 CCCCTTCCCACCCCTCTGCCGCCACCTTTCTTGGAGGCGACAACCCCGGGAGGCGATTAG 180  
QY 181 AAGGGATTTTCCCGCAGGTTGCGAAGGGAAGCAAACTTGTGGCAACTTGTCTCCCGGT 240  
DB 181 AAGGGATTTTCCCGCAGGTTGCGAAGGGAAGCAAACTTGTGGCAACTTGTCTCCCGGT 240  
QY 241 GCGGGCGTCTCTCCCCCACCCTGTCAACATGCTTAGGGGTCCGGGGCCGGGCTGTGCT 300  
DB 241 GCGGGCGTCTCTCCCCCACCCTGTCAACATGCTTAGGGGTCCGGGGCCGGGCTGTGCT 300  
QY 301 GCTGGCGCTCTGTGCTGGGACAGCGGTGCCCTCCACGGGAGGCTCGAAGAGCAAGAG 360  
DB 301 GCTGGCGCTCTGTGCTGGGACAGCGGTGCCCTCCACGGGAGGCTCGAAGAGCAAGAG 360  
QY 361 GCAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTGTCAGTCAAAAGCAAGCCCG 420  
DB 361 GCAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTGTCAGTCAAAAGCAAGCCCG 420  
QY 421 TTGTTATGCAATGGAAGAACATATCAGATAATCAACAGTGGGAGCGGCTTACCTAGG 480  
DB 421 TTGTTATGCAATGGAAGAACATATCAGATAATCAACAGTGGGAGCGGCTTACCTAGG 480  
QY 481 CAATGCGTTGTTGTTACTTGTATCGAGGAAACCGAGGTTTAACTGCGAAAGTAAACC 540  
DB 481 CAATGCGTTGTTGTTACTTGTATCGAGGAAACCGAGGTTTAACTGCGAGAGTAAACC 540  
QY 541 TGAAGCTGAAGAGACTTGTCTTTTGCAAGTACACTGCGGAACACTTACCAGTGGGTGACAC 600

541 TGAAGCTGAAGAGACTGTTGTCAGCAAGTACACTGGGACACCTTACCGAGTGGGTGACAC 600  
601 TTATGAGCGCTCTAAAGACTCCATGATCTGGGAGCTGTACTCTGCATCGGGCTGGGCGAGG 660  
601 TTATGAGCGCTCTAAAGACTCCATGATCTGGGAGCTGTACTCTGCATCTGGGGCTGGGCGAGG 660  
661 GAGAAATAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGGTCAAGTCTTACAAGATTGG 720  
661 GAGAAATAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGGTCAAGTCTTACAAGATTGG 720  
721 TGACACCTGGAGAGACCAATGAGACTGGTGTATACATGTTAGAGTGTGTGTCTGG 780  
721 TGACACCTGGAGAGACCAATGAGACTGGTGTATACATGTTAGAGTGTGTGTCTGG 780  
781 TAATGGAAGAGAGATGGACCTGCAAGCCATGAGCTGAGAGAGTGTGTTGATCATGCTGC 840  
781 TAATGGAAGAGAGATGGACCTGCAAGCCATGAGCTGAGAGAGTGTGTTGATCATGCTGC 840  
841 TGGGACTTCTATGTGGTCCGAGAGAAAGCTGGGAGAGCCCTACCAAGGCTGGATGATGGT 900  
841 TGGGACTTCTATGTGGTCCGAGAGAAAGCTGGGAGAGCCCTACCAAGGCTGGATGATGGT 900  
901 AGATTGTACTTGGCTGGGAGAGCGAGCGGACGCATCATCTGCACTTCTAGAAATAGATG 960  
901 AGATTGTACTTGGCTGGGAGAGCGAGCGGACGCATCATCTGCACTTCTAGAAATAGATG 960  
961 CAACGATCAGGACACCAAGGACATCTTATAGAAATGGAGACACCTGGAGCAAGAGGATAA 1020  
961 CAACGATCAGGACACCAAGGACATCTTATAGAAATGGAGACACCTGGAGCAAGAGGATAA 1020  
1021 TCGAGGAACTCTCTCAGTGCATCTGCACAGCAAGCGCGGAGAGTGGAGTGTGA 1080  
1021 TCGAGGAACTCTCTCAGTGCATCTGCACAGCAAGCGCGGAGAGTGGAGTGTGA 1080  
1081 GAGGACACCTCTGTGAGAGACCATCGAGCGGATCTGGCCCTTCAACGATCTTCTGTCG 1140  
1081 GAGGACACCTCTGTGAGAGACCATCGAGCGGATCTGGCCCTTCAACGATCTTCTGTCG 1140  
1141 AGCTGTTTACACCGAGGCTCAACCGGAGCTCTCCCTATGGGCACTGTGTACAGA 1200  
1141 AGCTGTTTACACCGAGGCTCAACCGGAGCTCTCCCTATGGGCACTGTGTACAGA 1200  
1201 CAGTGGTGTGGTCTACTCTGTGGGATGCTAGTGGCTGAAGACACCAAGGAAATAGCAAT 1260  
1201 CAGTGGTGTGGTCTACTCTGTGGGATGCTAGTGGCTGAAGACACCAAGGAAATAGCAAT 1260  
1261 GCTTTGCAAGTGGGCAACCGAGTCAAGAGACAGCTGTAAACCGAGACTTA 1320  
1261 GCTTTGCAAGTGGGCAACCGAGTCAAGAGACAGCTGTAAACCGAGACTTA 1320  
1321 CGGTGGCAACTCAATGGAGAGCCATGTGCTTACCAATCACTTCACTTCACTTCACTTCACT 1380  
1321 CGGTGGCAACTCAATGGAGAGCCATGTGCTTACCAATCACTTCACTTCACTTCACTTCACT 1380  
1381 CTACTCTCTGCAACCAAGAGGCGCAAGGAGCGGACATCTTTGTTGAGCAGCACTTCGAA 1440  
1381 CTACTCTCTGCAACCAAGAGGCGCAAGGAGCGGACATCTTTGTTGAGCAGCACTTCGAA 1440  
1441 TTATGAGCAGGACCAAGAAATCTCTTCTGACAGACCACTGTTTGGTTCAGACTCG 1500  
1441 TTATGAGCAGGACCAAGAAATCTCTTCTGACAGACCACTGTTTGGTTCAGACTCG 1500  
1501 AGGAGGAAATCCAAATGGTGGCTTGTGCGCACTTCCCTTCTTACCAACCAACCAATTA 1560  
1501 AGGAGGAAATCCAAATGGTGGCTTGTGCGCACTTCCCTTCTTACCAACCAACCAATTA 1560  
1561 CACTGATTGCACTTCTGAGGCAAGAGCAACATGAGTGGTGGGACCAACAGAA 1620  
1561 CACTGATTGCACTTCTGAGGCAAGAGCAACATGAGTGGTGGGACCAACAGAA 1620  
1621 CTATGATCCCGACCAAGATTGGGTTCTGCGCCATGGCTGCCACAGGAAATCTGAC 1680

Db 1621 CTATGATCCCGACAGAAAGTTTGGGTTCTGCCCATGGCTGCCACGAGGAAATCTGCAC 1680  
QY 1681 AACCAATGAGGGGTCTATGTACCGCATTTGGAGATCAGTGGGATAAGCAGCATGACATGG 1740  
Db 1681 AACCAATGAGGGGTCTATGTACCGCATTTGGAGATCAGTGGGATAAGCAGCATGACATGG 1740  
QY 1741 TCACATGATGAGTGCACGTGTGTTGGGAATGTCTGGGGAATGCAATGCAATTCGCTTA 1800  
Db 1741 TCACATGATGAGTGCACGTGTGTTGGGAATGTCTGGGGAATGCAATGCAATTCGCTTA 1800  
QY 1801 CTGCGACCTTCGAGATCAGTGCATTTGATGATCATCTTCAATGTTGAAACGACATTT 1860  
Db 1801 CTGCGACCTTCGAGATCAGTGCATTTGATGATCATCTTCAATGTTGAAACGACATTT 1860  
QY 1861 CCACAAGCGTTCATGAGAGGGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920  
Db 1861 CCACAAGCGTTCATGAGAGGGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920  
QY 1921 CAGTGGAGTGTGATCCCGTTCGACCAATGCGAGGATTCAGAGCTGGGACGTTTATCA 1980  
Db 1921 CAGTGGAGTGTGATCCCGTTCGACCAATGCGAGGATTCAGAGCTGGGACGTTTATCA 1980  
QY 1981 AATTGGAGATTTCATGGGAGAGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
Db 1981 AATTGGAGATTTCATGGGAGAGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
QY 2041 CCGTGGCATTTGGGAGTGGCATTTGCGAACTTTTACAGACCTTCCAGCTCAGCTGCTGCTG 2100  
Db 2041 CCGTGGCATTTGGGAGTGGCATTTGCGAACTTTTACAGACCTTCCAGCTCAGCTGCTGCTG 2100  
QY 2101 TGTGGAAGTATTTATCACTGAGACTCCGAGTCCAGCCCACTCCACCCCTCCAGTGGAA 2160  
Db 2101 TGTGGAAGTATTTATCACTGAGACTCCGAGTCCAGCCCACTCCACCCCTCCAGTGGAA 2160  
QY 2161 TGCACCAAGCCATCTCAGATTTCCAAAGTACATTTCCAGGTGGAGCTTAAATTTCTGT 2220  
Db 2161 TGCACCAAGCCATCTCAGATTTCCAAAGTACATTTCCAGGTGGAGCTTAAATTTCTGT 2220  
QY 2221 AGGCGTTTGAAGGAGAGCTACCATACAGGCGCACTTAAACTCTTACACCATCAAGGCT 2280  
Db 2221 AGGCGTTTGAAGGAGAGCTACCATACAGGCGCACTTAAACTCTTACACCATCAAGGCT 2280  
QY 2281 GAAGCTGCTGTGTATACAGAGGCGCACTCATCAGCATTCAGCAGTACGCGCCACCAAGA 2340  
Db 2281 GAAGCTGCTGTGTATACAGAGGCGCACTCATCAGCATTCAGCAGTACGCGCCACCAAGA 2340  
QY 2341 AGTGAAGT 2400  
Db 2341 AGTGAAGT 2400  
QY 2401 GACGAGGAGAGAGCTCCCTTTCTCTTGTGGGCACTTCTGAAATCTGTGACCGGAAAT 2460  
Db 2401 GACGAGGAGAGAGCTCCCTTTCTCTTGTGGGCACTTCTGAAATCTGTGACCGGAAAT 2460  
QY 2461 CACGACGATGCTGT 2520  
Db 2461 CACGACGATGCTGT 2520  
QY 2521 GGTGGAATATGAGT 2580  
Db 2521 GGTGGAATATGAGT 2580  
QY 2581 AGCCTCTCTGTGAACATCCCTGACCTGCTTCTGCGGCAATACATTTGAAATGCTTA 2640  
Db 2581 AGCCTCTCTGTGAACATCCCTGACCTGCTTCTGCGGCAATACATTTGAAATGCTTA 2640  
QY 2641 TCAGATATCTGAGGATGGGAGCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2700  
Db 2641 TCAGATATCTGAGGATGGGAGCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2700  
QY 2701 TGTGCGGCTCTGACCGGCTGT 2760  
Db 2701 TGTGCGGCTCTGACCGGCTGT 2760



4921	CGATGTTCCAGACACACAGCATAGTGTCAAGTGGTGGCTTCAAGTTCCTCCCTGTTACTGG	6001	CTTGGCCACCAACACCAATTCCTTCTGTATCATGTCAGCGCCGACGTCGACGATTAC	6060
4981	TTACAGAGTAACACCACTCCCAAAATGGACACAGGACCAACAAAATCTAAATCTGCAGG	6061	CGGCTACATCATCAAGTATGAGAACCTTGGTCTCTCTCCAGAGAGTGGTCCCTCGGCC	6120
4981	TTACAGAGTAACACCACTCCCAAAATGGACACAGGACCAACAAAATCTAAATCTGCAGG	6061	CGGCTACATCATCAAGTATGAGAACCTTGGTCTCTCTCCAGAGAGTGGTCCCTCGGCC	6120
5041	TCCAGATCAAAACAGAAATGACTATTGAAGGCTTGCAGCCACAGTGGAGTATGGTTAG	6121	CGCCCTGTGTACACAGAGGCTTACTATTACTGGCTTGAACCGGGAAACCGAATATACAAAT	6180
5041	TCCAGATCAAAACAGAAATGACTATTGAAGGCTTGCAGCCACAGTGGAGTATGGTTAG	6121	CGCCCTGTGTACACAGAGGCTTACTATTACTGGCTTGAACCGGGAAACCGAATATACAAAT	6180
5101	TGTCTATGCTCAGAAATCCAAAGCGGAGAGTCAAGCTTGGTTCAGAGCTGCAAGTAAACAA	6181	TTATGTCTATGTCCTTGAAGATTAATCAGAGAGCGAGCCCTGATTTGAAGGAAAGAAC	6240
5101	TGTCTATGCTCAGAAATCCAAAGCGGAGAGTCAAGCTTGGTTCAGAGCTGCAAGTAAACAA	6181	TTATGTCTATGTCCTTGAAGATTAATCAGAGAGCGAGCCCTGATTTGAAGGAAAGAAC	6240
5161	CATTGATCGCCCTAAAGACTTGGCATTGATGTGATGTGATGTGATTCATCAAAATGTC	6241	AGACGAGCTTCCCAACCTTCCACACCCCAATCTTCAATGACACGAGATCTT	6300
5161	CATTGATCGCCCTAAAGACTTGGCATTGATGTGATGTGATGTGATTCATCAAAATGTC	6241	AGACGAGCTTCCCAACCTTCCACACCCCAATCTTCAATGACACGAGATCTT	6300
5221	TTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTCGAGCCCTGAGGA	6301	GGATGTTCTTCCACAGTTCAAAAGACCCCTTTCGTACCCACCCCTGGGTATGACACTGG	6360
5221	TTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTCGAGCCCTGAGGA	6301	GGATGTTCTTCCACAGTTCAAAAGACCCCTTTCGTACCCACCCCTGGGTATGACACTGG	6360
5281	TGGAAATCCATGAGCTATTCCCTGCACTGATGTGATGTGATGTGATGTGATGTGATGTG	6361	AAATGTATTCAGCTTCCCTGGCACTTCTGTCAGCAACCCAGTGTGGGCAACAATGAT	6420
5281	TGGAAATCCATGAGCTATTCCCTGCACTGATGTGATGTGATGTGATGTGATGTGATGTG	6361	AAATGTATTCAGCTTCCCTGGCACTTCTGTCAGCAACCCAGTGTGGGCAACAATGAT	6420
5341	CCTCAGACCGGGTTCAGATACACAGTCAAGTGTGATGTGATGTGATGTGATGTGATGTG	6421	CTTTGAGGAACATGTTTAAAGCGGACCAACCGGCCCAACCGGCCACCCCATTAAGGCA	6480
5341	CCTCAGACCGGGTTCAGATACACAGTCAAGTGTGATGTGATGTGATGTGATGTGATGTG	6421	CTTTGAGGAACATGTTTAAAGCGGACCAACCGGCCCAACCGGCCACCCCATTAAGGCA	6480
5401	CCAGCCCTGATTTGAAACCCAGTCCACAGCTATTCCTGCAACAACTGACCTGAAAGTTTCA	6481	TAGGCCAGACCATACCCGCGAATGTAGGACAAAGAGTCTCTCAGACAAACATCTC	6540
5401	CCAGCCCTGATTTGAAACCCAGTCCACAGCTATTCCTGCAACAACTGACCTGAAAGTTTCA	6481	TAGGCCAGACCATACCCGCGAATGTAGGACAAAGAGTCTCTCAGACAAACATCTC	6540
5461	TCAGTCAACCCCAAGCCTGAGCGCCAGTGGACACACCCATGTTGACCTGACCTGAAAGTT	6541	ATGGGCCCATTCACAGGACACTTCTGAGTACATCAATTCATGTCTCTCTGTTGGCACTGA	6600
5461	TCAGTCAACCCCAAGCCTGAGCGCCAGTGGACACACCCATGTTGACCTGACCTGAAAGTT	6541	ATGGGCCCATTCACAGGACACTTCTGAGTACATCAATTCATGTCTCTCTGTTGGCACTGA	6600
5521	ATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAATCAACCTTGC	6601	TGAAGAACCTTACAGTTCAGGGTTCCTGGAACTTCTACAGTGGCACTCTGACAGGCT	6660
5521	ATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAATCAACCTTGC	6601	TGAAGAACCTTACAGTTCAGGGTTCCTGGAACTTCTACAGTGGCACTCTGACAGGCT	6660
5581	TCCTCAGAGCTCATCCGTTGTTATCAGGACTTATGTTGGCCCAACAAATATGAAGTGAG	6661	CACAGAGTGGCCACCTACCAACATCATAGTGGAGGACCTGAAAGACCCAGAGAGCATAA	6720
5581	TCCTCAGAGCTCATCCGTTGTTATCAGGACTTATGTTGGCCCAACAAATATGAAGTGAG	6661	CACAGAGTGGCCACCTACCAACATCATAGTGGAGGACCTGAAAGACCCAGAGAGCATAA	6720
5641	TGTCTATGCTTTAAGGACACTTTGCAAGCAGACAGCTCAGGGAGTGTGCAACACTCT	6721	GGTTCGGGAGAGGTTGTTACCGTGGGCACTCTGTCAACGAAAGCTTGAACCAACCTAC	6780
5641	TGTCTATGCTTTAAGGACACTTTGCAAGCAGACAGCTCAGGGAGTGTGCAACACTCT	6721	GGTTCGGGAGAGGTTGTTACCGTGGGCACTCTGTCAACGAAAGCTTGAACCAACCTAC	6780
5701	GGAGAAATGTCAGCCCAAGAGGGCTGTTGTGACAGATGCTACTGAGACCAATCAC	6781	GGATGACTGCTGCTTTGACCCCTTACACAGTGTCCCATATATGCGTTGGAGATGAGTGGGA	6840
5701	GGAGAAATGTCAGCCCAAGAGGGCTGTTGTGACAGATGCTACTGAGACCAATCAC	6781	GGATGACTGCTGCTTTGACCCCTTACACAGTGTCCCATATATGCGTTGGAGATGAGTGGGA	6840
5761	CATTAGCTGGAGAAACCAAGACTGAGACGATCACTGGCTTCCAGTTGATGCGGTTCCAGC	6841	ACGAAATGTCTGAATCAGGCTTTAAACGTTGTGCGACGCTTAGGCTTTGGAAGTGGTCA	6900
5761	CATTAGCTGGAGAAACCAAGACTGAGACGATCACTGGCTTCCAGTTGATGCGGTTCCAGC	6841	ACGAAATGTCTGAATCAGGCTTTAAACGTTGTGCGACGCTTAGGCTTTGGAAGTGGTCA	6900
5821	CAATGGCCAGATCCAAATCCAGAGAACCAATGAGCCAGATGTGAGAGCTACCAATCAC	6901	TTTCAGATGTGATTCATCTAGATGGTCCATGACAAATGTTGTAACCAAGATTTGGAGA	6960
5821	CAATGGCCAGATCCAAATCCAGAGAACCAATGAGCCAGATGTGAGAGCTACCAATCAC	6901	TTTCAGATGTGATTCATCTAGATGGTCCATGACAAATGTTGTAACCAAGATTTGGAGA	6960
5881	TGCTTTACAAACAGGACCTGACTACAGATCTACCTGTACACTTGAATGCAATGCTCG	6961	GAAAGTGGGACCGTCAGGAGAAATGGCCAGATGATGAGCTGCATGCTTTGGGAAACGG	7020
5881	TGCTTTACAAACAGGACCTGACTACAGATCTACCTGTACACTTGAATGCAATGCTCG	6961	GAAAGTGGGACCGTCAGGAGAAATGGCCAGATGATGAGCTGCATGCTTTGGGAAACGG	7020
5941	GAGTCCCTGTGTCATCGACGCTTCCATGCTGATGATGATGATGATGATGATGATGATG	7021	AAAGGAGGAAATCAAGTGTGACCTCATGAGGCAACGCTGTACGATGATGGGAAACATA	7080
5941	GAGTCCCTGTGTCATCGACGCTTCCATGCTGATGATGATGATGATGATGATGATGATG	7021	AAAGGAGGAAATCAAGTGTGACCTCATGAGGCAACGCTGTACGATGATGGGAAACATA	7080
6001	CTTGGCCACCAACCAATTCCTTCTGTATCATGTCAGCGCCGACGTCGACGATTAC	7081	CCAGTAGGAGAACAGTGGCAGAGGAAATATCTCGGTGCCATTTGCTCTGCAATGCTTT	7140
6001	CTTGGCCACCAACCAATTCCTTCTGTATCATGTCAGCGCCGACGTCGACGATTAC	7081	CCAGTAGGAGAACAGTGGCAGAGGAAATATCTCGGTGCCATTTGCTCTGCAATGCTTT	7140

```

7141 TGAAGCCAGCGGGCTGGCGCTGTGCAACTCCGCGAGACCTGGGGTGAACCCAGTCC 7200
7141 TGAAGCCAGCGGGCTGGCGCTGTGCAACTCCGCGAGACCTGGGGTGAACCCAGTCC 7200
7201 CGAAGCACTACTGGCCAGTCTCAACACCAAGTATTTCTCAGAGATACCATCAGAGACAAA 7260
7201 CGAAGCACTACTGGCCAGTCTCAACACCAAGTATTTCTCAGAGATACCATCAGAGACAAA 7260
7261 CACTAATGTTAATGGCCCAATGAGTGTCTCATGCTTTCATGCTTTAGATGTACAGCTGACAGAGA 7320
7261 CACTAATGTTAATGGCCCAATGAGTGTCTCATGCTTTCATGCTTTAGATGTACAGCTGACAGAGA 7320
7321 AGATTCGCGAGAGTAATCATCTTTCCAAATCCAGAGGAAACAAGCATGTCTCTGCGCAAG 7380
7321 AGATTCGCGAGAGTAATCATCTTTCCAAATCCAGAGGAAACAAGCATGTCTCTGCGCAAG 7380
7381 ATCCATCTAACTGGAGTGATGTTAGCAGACCCAGCTTAGAGTTCTTCTTTCTTTAA 7440
7381 ATCCATCTAACTGGAGTGATGTTAGCAGACCCAGCTTAGAGTTCTTCTTTCTTTAA 7440
7441 GCCCTTTGCTCTGAGGAAGTCTCCAGCTTCAGCTCAACTCAGAGCTTCTCCAGCATC 7500
7441 GCCCTTTGCTCTGAGGAAGTCTCCAGCTTCAGCTCAACTCAGAGCTTCTCCAGCATC 7500
7501 ACCCTGGAGTTTCTGAGGGTTTCTCAATAATGAGGGCTGCACATGCTGCTTCTGCT 7560
7501 ACCCTGGAGTTTCTGAGGGTTTCTCAATAATGAGGGCTGCACATGCTGCTTCTGCT 7560
7561 TCGAAGTATTCATACCGCTCAGTATTTTAAATGAAGTATTTCTAAGATTTGGTTTGGGA 7620
7561 TCGAAGTATTCATACCGCTCAGTATTTTAAATGAAGTATTTCTAAGATTTGGTTTGGGA 7620
7621 TCAATAGGAAGCATATGAGCCCAACCAAGATGCAAAATGTTTGAATGATATGACCAAA 7680
7621 TCAATAGGAAGCATATGAGCCCAACCAAGATGCAAAATGTTTGAATGATATGACCAAA 7680
7681 ATTTTAGTAGGAAGTACCCAAACACTTCTGCTTCACTTAAGTGTCTGCGCCCGCAAT 7740
7681 ATTTTAGTAGGAAGTACCCAAACACTTCTGCTTCACTTAAGTGTCTGCGCCCGCAAT 7740
7741 ACTGTAGGAACCAAGCATGATCTTGTACTGTGATATTTAAATATCCACAGTACT 7795
7741 ACTGTAGGAACCAAGCATGATCTTGTACTGTGATATTTTAAATATCCACAGTACT 7795

```

```

RESULT 3
US-10-098-841-8
; Sequence 8, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyen
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; FILE REFERENCE: 784C12
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042

```

```

; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 8
; LENGTH: 8137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (363)..(7430)
US-10-098-841-8

Query Match 99.7%; Score 7774.8; DB 13; Length 8137;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 7779; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 10 CCGGCTGTGTCACAGGGGAGGAGGGAAACCCAGGCGGCGAGCGGGAAGAGGGGACC 69
DB 104 CCGGCTGTGTCACAGGGGAGGAGGGAAACCCAGGCGGCGAGCGGGAAGAGGGGACC 163
QY 70 TGCAGCCACAATCTTCTGTCCTCTGTCATCCCTTCTGTCCTCCACCCGTCCTCCCTCC 129
DB 164 TGCAGCCACAATCTTCTGTCCTCTGTCATCCCTTCTGTCCTCCACCCGTCCTCCCTCC 223
QY 130 CACCTCTGGGCCCCACCTTCTTGGAGGCGACAACCCCGGAGGCGATTAGAAGGGATT 189
DB 224 CACCTCTGGGCCCCACCTTCTTGGAGGCGACAACCCCGGAGGCGATTAGAAGGGATT 283
QY 190 TTCCCGCAGGTTGGAGGGAAGCAAACTTGTGCGCAACTTGTCTCCCTCCGCTGCGGGGTC 249
DB 284 TTCCCGCAGGTTGGAGGGAAGCAAACTTGTGCGCAACTTGTCTCCCTCCGCTGCGGGGTC 343
QY 250 TCTCCCCCACCCTCTCAACATGCTTAGGGGTCCGGGGCCCGGGCTGTGCTGCTGCGCGGT 309
DB 344 TCTCCCCCACCCTCTCAACATGCTTAGGGGTCCGGGGCCCGGGCTGTGCTGCTGCGCGGT 403
QY 310 CTGTGCTCTGGGAGCAGGGTCCCTCCACGGAGCCTCGAAGAGCAAGAGGAGGCTCA 369
DB 404 CTGTGCTCTGGGAGCAGGGTCCCTCCACGGAGCCTCGAAGAGCAAGAGGAGGCTCA 463
QY 370 GCAAAATGTTTCAGCCCGCAGTCCCGGTGCTGTCAGTCAAAAGCAAGCCGGTGTGTTATGA 429
DB 464 GCAAAATGTTTCAGCCCGCAGTCCCGGTGCTGTCAGTCAAAAGCAAGCCGGTGTGTTATGA 523
QY 430 CAATCGAAAAACATATCAGATAAAATCAACAGTGGAGCGGACCTACCTAGGCAATGCGTT 489
DB 524 CAATCGAAAAACATATCAGATAAAATCAACAGTGGAGCGGACCTACCTAGGCAATGCGTT 583
QY 490 GGTGTTGTTACTTGTATGAGGAGGAGCCGAGGTTTAACTGCGAAAAGTAACTGAGGCTGA 549
DB 584 GGTGTTGTTACTTGTATGAGGAGGAGCCGAGGTTTAACTGCGAGAAGTAACTGAGGCTGA 643
QY 550 AGAGACTTGTCTTTGACAAAGTACACTTGGGAAACACTTACCGAGTGGGTGACACTTATGAGCG 609
DB 644 AGAGACTTGTCTTTGACAAAGTACACTTGGGAAACACTTACCGAGTGGGTGACACTTATGAGCG 703
QY 610 TCCTAAAGACTCCATGATCTGGGACTGTACTGTCATCGGGCTGGGCGGAGGAGGAAATAG 669
DB 704 TCCTAAAGACTCCATGATCTGGGACTGTACTGTCATCGGGCTGGGCGGAGGAGGAAATAG 763
QY 670 CTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCAAGTCTTACAAGATTGGTGACACCTG 729
DB 764 CTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCAAGTCTTACAAGATTGGTGACACCTG 823
QY 730 GAGGAGACCATGAGACTGTTGTTTACATGTTAGAGTGTGTGTCTTCTGTAATGGA 789
DB 824 GAGGAGACCATGAGACTGTTGTTTACATGTTAGAGTGTGTGTCTTCTGTAATGGA 883
QY 790 AGGAGAAATGAGACCTGCAAGCCCATAGCTGAGAAAGTGTGTTTGTATCATGCTGCTGGGACTTC 849

```







QY	3010	GTGTTGGAAGTGCAGACGCTGAAGGTCAACCATGTGGACACCGCTCAGAGTGCAGT	3069	QY	4090	ATTACCAACATTGTTCCAGACACCATCGGTGTCACTGGGCTCCACCCCATTCATGA	4149
DB	3104	GTGTTGGAAGTGCAGACGCTGAAGGTCAACCATGTGGACACCGCTCAGAGTGCAGT	3163	DB	4184	ATTCACCAACATTGTTCCAGACACCATCGGTGTCACTGGGCTCCACCCCATTCATGA	4243
QY	3070	GACCGGCTACCGTGTGATGTGATCCCGGTCAACCTGCTGGCGAGCAGCGGAGAGGCT	3129	QY	4150	TTTAAACCAACTTCCTGGTGCCTTACTCACCTGTGAAAAATGAGGAAGATGTTGCAGAGTT	4209
DB	3164	GACCGGCTACCGTGTGATGTGATCCCGGTCAACCTGCTGGCGAGCAGCGGAGAGGCT	3223	DB	4244	TTTAAACCAACTTCCTGGTGCCTTACTCACCTGTGAAAAATGAGGAAGATGTTGCAGAGTT	4303
QY	3130	GCCCATCAGCAGGAACAACCTTTGAGAGAGTCAACCGGCTGTCCCTGGGGTCACTATTATTA	3189	QY	4210	GTCAATTTCTCCTTCAGACAAATGCGAGTGGTCTTAACAAATCTCCTGCCCTGGTACAGATA	4269
DB	3224	GCCCATCAGCAGGAACAACCTTTGAGAGAGTCAACCGGCTGTCCCTGGGGTCACTATTATTA	3283	DB	4304	GTCAATTTCTCCTTCAGACAAATGCGAGTGGTCTTAACAAATCTCCTGGCTGGTACAGATA	4363
QY	3190	CTTCAAAGTCTTTGAGTGAAGCCATGGGAGGAGAGCAGAGCTCTGACTCTCAACAGAC	3249	QY	4270	TGTAGTGAAGTGTCTCCAGTGTCTACGAAACAACATGAGAGCACACCTCTTAGAGGAAGACA	4329
DB	3284	CTTCAAAGTCTTTGAGTGAAGCCATGGGAGGAGAGCAGAGCTCTGACTCTCAACAGAC	3343	DB	4364	TGTAGTGAAGTGTCTCCAGTGTCTACGAAACAACATGAGAGCACACCTCTTAGAGGAAGACA	4423
QY	3250	AACCAAACTGGATGCTCCCATTAACCTCCAGTTTGTCAATGAAGTCAATCTACTGTCTCT	3309	QY	4330	GAAAACAGGCTCTTGATTTCCCAACTGGGATTTGACCTTTCTGATATTACTGCCAATCTTTT	4389
DB	3344	AACCAAACTGGATGCTCCCATTAACCTCCAGTTTGTCAATGAAGTCAATCTACTGTCTCT	3403	DB	4424	GAAAACAGGCTCTTGATTTCCCAACTGGGATTTGACCTTTCTGATATTACTGCCAATCTTTT	4483
QY	3310	GGTGAAGTGAATCTCACCTCGGGGCCAGATTAACAGGATACCGACTGACCGTGGGCCCTTAC	3369	QY	4390	TACTGTGACTGGATTTGTCTCTCGAGGCCACCATCACTGGCTACAGGATCCGCCATCATCC	4449
DB	3404	GGTGAAGTGAATCTCACCTCGGGGCCAGATTAACAGGATACCGACTGACCGTGGGCCCTTAC	3463	DB	4484	TACTGTGACTGGATTTGTCTCTCGAGGCCACCATCACTGGCTACAGGATCCGCCATCATCC	4543
QY	3370	CCGAGAGGACAGCCAGGAGTCAATGTGGGTCCCTCTGTCTCCAAAGTACCCACTGAG	3429	QY	4450	CGAGCACTTCAGTGGGAGACCTCGAGAAAGATCGGGTGGCCCACTCTCGGAATTCATCACC	4509
DB	3464	CCGAGAGGACAGCCAGGAGTCAATGTGGGTCCCTCTGTCTCCAAAGTACCCACTGAG	3523	DB	4544	CGAGCACTTCAGTGGGAGACCTCGAGAAAGATCGGGTGGCCCACTCTCGGAATTCATCACC	4603
QY	3430	GAATCTGAGCTGATCTGAGTACACCGTATTCCTGTGGTCCCATAAAGGGCAACCAAGA	3489	QY	4510	CCTCACCACCTCACTCCAGGACACAGAGTATGTGTGTCAGCATCGTTGCTTTAAATGGCAG	4569
DB	3524	GAATCTGAGCTGATCTGAGTACACCGTATTCCTGTGGTCCCATAAAGGGCAACCAAGA	3583	DB	4604	CCTCACCACCTCACTCCAGGACACAGAGTATGTGTGTCAGCATCGTTGCTTTAAATGGCAG	4663
QY	3490	GAGCCCCAAGCCATGAGTCTTTACCACATGCGAGCTCGGAGCTCTATTCCACCTTA	3549	QY	4570	AGAGGAAATCCCTTTATTGATGGCCAAACAATCAACAGTTTCTGATGTTCCGAGGAGACT	4629
DB	3584	GAGCCCCAAGCCATGAGTCTTTACCACATGCGAGCTCGGAGCTCTATTCCACCTTA	3643	DB	4664	AGAGGAAATCCCTTTATTGATGGCCAAACAATCAACAGTTTCTGATGTTCCGAGGAGACT	4723
QY	3550	CAACACCGAGTGAATGAGACACCATTTGATCACATGGAAGCTGCTCCAAAGATTGG	3609	QY	4630	GGAGTGTGTTGTCGACCCCAACCGCTACTGATCAGCTGGGATGCTCCTGCTGTAC	4689
DB	3644	CAACACCGAGTGAATGAGACACCATTTGATCACATGGAAGCTGCTCCAAAGATTGG	3703	DB	4724	GGAGTGTGTTGTCGACCCCAACCGCTACTGATCAGCTGGGATGCTCCTGCTGTAC	4783
QY	3610	TTTTAAGTGGGTGACACCAAGCCAGGAGGAGGACCAACGAGAGTGAATTCAGA	3669	QY	4690	AGTGAGATATTACAGGATCACTTACCGGAGAAACAGAGGAAATAGCCCTGTCCAGAGTT	4749
DB	3704	TTTTAAGTGGGTGACACCAAGCCAGGAGGAGGACCAACGAGAGTGAATTCAGA	3763	DB	4784	AGTGAGATATTACAGGATCACTTACCGGAGAAACAGAGGAAATAGCCCTGTCCAGAGTT	4843
QY	3670	CTCAGGAAGCATCGTTGTGTCGGCTTGACCTCCAGGAGTAGAATACGTTACACCATCCA	3729	QY	4750	CACTGTGCTGGAGCAAGTCTACAGCTACCATCAGCGCCTTAAACCTGGAGTTGATTA	4809
DB	3764	CTCAGGAAGCATCGTTGTGTCGGCTTGACCTCCAGGAGTAGAATACGTTACACCATCCA	3823	DB	4844	CACTGTGCTGGAGCAAGTCTACAGCTACCATCAGCGCCTTAAACCTGGAGTTGATTA	4903
QY	3730	AGTCTGAGAGTGAAGAGAGATGCGGCAATTGTAAACAAAGTGGTGAACCATTT	3789	QY	4810	TACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAGCCCGCAAGCAGCAAGCCAAAT	4869
DB	3824	AGTCTGAGAGTGAAGAGAGATGCGGCAATTGTAAACAAAGTGGTGAACCATTT	3883	DB	4904	TACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAGCCCGCAAGCAGCAAGCCAAAT	4963
QY	3790	GTCTCCACCAACAACTTGCATCTGGAGCAAAACCTGCATCTGAGTGTCTCACAGTCTC	3849	QY	4870	TTCCATTAATTACCGAACAAGAAATTGAACCAATCCCGAGATGCAAGTGACCGATGTTCA	4929
DB	3884	GTCTCCACCAACAACTTGCATCTGGAGCAAAACCTGCATCTGAGTGTCTCACAGTCTC	3943	DB	4964	TTCCATTAATTACCGAACAAGAAATTGAACCAATCCCGAGATGCAAGTGACCGATGTTCA	5023
QY	3850	CTGGAGAGGAGCAACACCCAGACATTACTGTTATAGAAATTACCAACACCCCTACAAA	3909	QY	4930	GGACAAACAGATTAGTGTCAAGTGGCTGCTTCAAGTTCCCTGTACTGGTTACAGAGT	4989
DB	3944	CTGGAGAGGAGCAACACCCAGACATTACTGTTATAGAAATTACCAACACCCCTACAAA	4003	DB	5024	GGACAAACAGATTAGTGTCAAGTGGCTGCTTCAAGTTCCCTGTACTGGTTACAGAGT	5083
QY	3910	CGGCCAGCAGGGAATTTCTTGGAGAGAGTGGTCCATGCTGATCAGAGCTCTCTGACCTTT	3969	QY	4990	AACCAACCTCCCAAAAATGGACAGGACCAACAAAACCTAATACTGAGGTTCCAGATCA	5049
DB	4004	CGGCCAGCAGGGAATTTCTTGGAGAGAGTGGTCCATGCTGATCAGAGCTCTCTGACCTTT	4063	DB	5084	AACCAACCTCCCAAAAATGGACAGGACCAACAAAACCTAATACTGAGGTTCCAGATCA	5143
QY	3970	TGATAACCTGAGTCCCGGCTGGAGTAAATGTCTAGTGTGTTTACCTGTCAAGAGTACAA	4029	QY	5050	AACAGAAATGACTATTGAAAGGCTTGGAGCCCAAGTGGAGTATGTTGTTGTTCTATGC	5109
DB	4064	TGATAACCTGAGTCCCGGCTGGAGTAAATGTCTAGTGTGTTTACCTGTCAAGAGTACAA	4123	DB	5144	AACAGAAATGACTATTGAAAGGCTTGGAGCCCAAGTGGAGTATGTTGTTGTTCTATGC	5203
QY	4030	GGAAAGTGTCCCTACTCTGATACCATCATCCAGCTGCTCCTCCCTCCACTGACCTGCG	4089	QY	5110	TCAGAAATCCAAGCGGAGAGAGTCAAGCTCTGTTTCAAGTGTGAGTAAACCAACATTGATCG	5169
DB	4124	GGAAAGTGTCCCTACTCTGATACCATCATCCAGCTGCTCCTCCCTCCACTGACCTGCG	4183	DB	5204	TCAGAAATCCAAGCGGAGAGAGTCAAGCTCTGTTTCAAGTGTGAGTAAACCAACATTGATCG	5263
				QY	5170	CCCTAAAGGACTGGCAATTCATGATGTGATGCTGATTCATCAAAATTCCTTGGGAAAG	5229

b CCTTAAAGAGCTGGCATTCACCTGATGTGATGTGATTCATCAAAATTCCTTGGGAAG 5323  
 y CCCACAGGGGCAAGTTTCAGGTACAGGTGAGTACCTACCTGAGCCCTGAGGATGGAATCCA 5289  
 b CCCACAGGGGCAAGTTTCAGGTACAGGTGAGTACCTACCTGAGCCCTGAGGATGGAATCCA 5383  
 y TGAGCTATTCCCTGCACTGATGTTGGAAGACACCTGAGAGCTGCAAGGCTCAGAGC 5349  
 b TGAGCTATTCCCTGCACTGATGTTGGAAGACACCTGAGAGCTGCAAGGCTCAGAGC 5443  
 y GGGTTCAGTACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5409  
 b GGGTTCAGTACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5503  
 y GATTGGAACCCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 5469  
 b GATTGGAACCCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 5563  
 y ACCACAAGCCTGAGGCGCCAGTGGACACCAACCAATGTTCAAGTCAAGTCAAGTCAAGTCA 5529  
 b ACCACAAGCCTGAGGCGCCAGTGGACACCAACCAATGTTCAAGTCAAGTCAAGTCAAGTCA 5623  
 y GGGGTGACCCCAAGGAGACCGGACCAATGAAGAAATCAACCTTGCTCCTGACAG 5589  
 b GGGGTGACCCCAAGGAGACCGGACCAATGAAGAAATCAACCTTGCTCCTGACAG 5683  
 y CTCATCCGTGGTGTATCAGGACTTATGTTGGGCCACCAAAATGATGAAGTGAAGTGTCTATGC 5649  
 b CTCATCCGTGGTGTATCAGGACTTATGTTGGGCCACCAAAATGATGAAGTGAAGTGTCTATGC 5743  
 y TCTTAAGGACACTTTCACAGCAGACCTCAGGAGTGTCAACCTCTGGGAATGT 5709  
 b TCTTAAGGACACTTTCACAGCAGACCTCAGGAGTGTCAACCTCTGGGAATGT 5803  
 y CAGCCACCAAGAGGCTCGTGTGACAGATGCTACTGAGACCAACCATCACCATTAGCTG 5769  
 b CAGCCACCAAGAGGCTCGTGTGACAGATGCTACTGAGACCAACCATCACCATTAGCTG 5863  
 y GAGAACCAAGACTGAGACCTCAGGAGTGTCAACCTCTGGGAATGT 5829  
 b GAGAACCAAGACTGAGACCTCAGGAGTGTCAACCTCTGGGAATGT 5923  
 y GACTCCAAATCCAGAGAACCATCAAGCAGATGTCAAGAGCTACACCATCAAGTGTACA 5889  
 b GACTCCAAATCCAGAGAACCATCAAGCAGATGTCAAGAGCTACACCATCAAGTGTACA 5983  
 y ACCAGGACCTGACTCAAGATCTACCTGACACCTTGAATGACAAATGCTCGAGCTCCCC 5949  
 b ACCAGGACCTGACTCAAGATCTACCTGACACCTTGAATGACAAATGCTCGAGCTCCCC 6043  
 y TGTGGTCACTCGAGCCTTCCACTGCTGATGATGATGATGATGATGATGATGATGATGATG 6009  
 b TGTGGTCACTCGAGCCTTCCACTGCTGATGATGATGATGATGATGATGATGATGATGATG 6103  
 y CACACCAATTCCTTGTGTATCATGGCAGCGCCACCTGCGAGATTAACCGGCTACAT 6069  
 b CACACCAATTCCTTGTGTATCATGGCAGCGCCACCTGCGAGATTAACCGGCTACAT 6163  
 y CATCAAGTATGAGAGCTCGGCTCTCTCCAGAGAGTGTCTCTCCGCGCCCGCCCTGG 6129  
 b CATCAAGTATGAGAGCTCGGCTCTCTCCAGAGAGTGTCTCTCCGCGCCCGCCCTGG 6223  
 y TGTCACAGAGGCTACTATTCCTGGCTGGAACCGGGAAACCGAATATACAAATTTATGTCT 6189  
 b TGTCACAGAGGCTACTATTCCTGGCTGGAACCGGGAAACCGAATATACAAATTTATGTCT 6283  
 y TGCCCTGAAGAAATATCAGAGAGCGAGCCCTGATTTGGAAGGAAAAGACAGAGAGCT 6249  
 b TGCCCTGAAGAAATATCAGAGAGCGAGCCCTGATTTGGAAGGAAAAGACAGAGAGCT 6343  
 y TCCCAACTGGTAAACCTTCCACACCCCAATCTTCATGGACCAAGAGATCTTGGATTTCC 6309

Db TCCCAACTGGTAAACCTTCCACACCCCAATCTTCATGACCAAGAGATCTTGGATGTTC 6403  
 Qy TCCCAAGTTCAAAAGACCCCTTTCCTACCCACCCCTGGGTATGACACTGGAAATGGTAT 6369  
 Db TCCCAAGTTCAAAAGACCCCTTTCCTACCCACCCCTGGGTATGACACTGGAAATGGTAT 6463  
 Qy TCAGCTTCCTGGCACTTCCTGGTCAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGA 6429  
 Db TCAGCTTCCTGGCACTTCCTGGTCAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGA 6523  
 Qy ACATGGTTTTAGCGGACACACCGCCCAACCGCCACCCCATAGGCATAGGCAAG 6489  
 Db ACATGGTTTTAGCGGACACACCGCCCAACCGCCACCCCATAGGCAATAGGCAAG 6583  
 Qy ACCATAACCCCGCAATGTAGGACAAAGAGCTCTCTCTGACAAACCATCTCATGGGCCCC 6549  
 Db ACCATAACCCCGCAATGTAGGACAAAGAGCTCTCTCTGACAAACCATCTCATGGGCCCC 6643  
 Qy ATTCCAGGACACTTCTGAGTACATCATTTTCATGTCATCTCTGTCGACCTGATGAAGAAC 6609  
 Db ATTCCAGGACACTTCTGAGTACATCATTTTCATGTCATCTCTGTCGACCTGATGAAGAAC 6703  
 Qy CTTCAGTTCAGGGTTCCTGAAACTTCTACCACTGCACTCTGACAGGCTCACCAGAG 6669  
 Db CTTCAGTTCAGGGTTCCTGAAACTTCTACCACTGCACTCTGACAGGCTCACCAGAG 6763  
 Qy TGCCACCTACAAATCATAGTGGAGCACTGAAGACCAAGACAGAGGATAGGTTGGGA 6729  
 Db TGCCACCTACAAATCATAGTGGAGCACTGAAGACCAAGACAGAGGATAGGTTGGGA 6823  
 Qy AGAGTTTGTACCGTGGCCAACTCTGTCAACGAAGCTTGAACCAACCTACCGATGACTC 6789  
 Db AGAGTTTGTACCGTGGCCAACTCTGTCAACGAAGCTTGAACCAACCTACCGATGACTC 6883  
 Qy GTGTTTGAACCCCTACACAGTGTCCCATTTAGCGTTGGAGATGAGTGGAAACGATGTC 6849  
 Db GTGTTTGAACCCCTACACAGTGTCCCATTTAGCGTTGGAGATGAGTGGAAACGATGTC 6943  
 Qy TGAATCAGGCTTTAACTGTTGCGCAGTCTTGAAGTGGTTCATTTAGATG 6909  
 Db TGAATCAGGCTTTAACTGTTGCGCAGTCTTGAAGTGGTTCATTTAGATG 7003  
 Qy TGATTCATCTAGATGTGTCATGACAAATGTGTGAACCTACAAATTTGGAGTGGGA 6969  
 Db TGATTCATCTAGATGTGTCATGACAAATGTGTGAACCTACAAATTTGGAGTGGGA 7063  
 Qy CCGTCAGGAGAAATGGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 7029  
 Db CCGTCAGGAGAAATGGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 7123  
 Qy ATTCAAGTGTGACCTCATGAGCAACCTGTTACGATGATGATGATGATGATGATGATGATG 7089  
 Db ATTCAAGTGTGACCTCATGAGCAACCTGTTACGATGATGATGATGATGATGATGATGATG 7183  
 Qy AGACAGTGGCAGAGAAATATCTCGTGCCATTTGCTCTGACATGCTTGGAGGCA 7149  
 Db AGACAGTGGCAGAGAAATATCTCGTGCCATTTGCTCTGACATGCTTGGAGGCA 7243  
 Qy CGGGGCTGGCGCTGTGACAACTGCGCAGACCTGGGGGTGAACCCAGTCCCGAAGCAC 7209  
 Db CGGGGCTGGCGCTGTGACAACTGCGCAGACCTGGGGGTGAACCCAGTCCCGAAGCAC 7303  
 Qy TACTGGCCGCTCTACAAACAGATTTCTCAGAGATACCATCAGAGACCAACATATGT 7269  
 Db TACTGGCCGCTCTACAAACAGATTTCTCAGAGATACCATCAGAGACCAACATATGT 7363  
 Qy TAATTTGCCAATTCAGTGTCTTCATGCTTTAGATGTACAGGCTGACAGAGAGATTTCCCG 7329  
 Db TAATTTGCCAATTCAGTGTCTTCATGCTTTAGATGTACAGGCTGACAGAGAGATTTCCCG 7423  
 Qy AGAGTAAATCATCTTTCCATCCAGAGAACAGATGTCTCTCTGCAAGATCCATCTA 7389  
 Db AGAGTAAATCATCTTTCCATCCAGAGAACAGATGTCTCTCTGCAAGATCCATCTA 7483

```

2y 7390 AACTGGAGTGAATTTAGCAGACCCAGGCTTAGAGTTCTTTCTTTCTTTAAAGCCCTTTGC 7449
2b 7484 AACTGGAGTGAATTTAGCAGACCCAGGCTTAGAGTTCTTTCTTTCTTTAAAGCCCTTTGC 7543
2y 7450 TCTGGAGGAAGTTCTCCAGCTTCAAGTCAACTCAGAGTTCTTCAAGAGTCAAGCTGGGA 7509
2b 7544 TCTGGAGGAAGTTCTCCAGCTTCAAGTCAACTCAGAGTTCTTCAAGAGTCAAGCTGGGA 7603
2y 7510 GTTTCCTGAGGTTTCTCTATAAATAGAGGCTGACATTCCTGTTCTGTTTCAAGATAT 7569
2b 7604 GTTTCCTGAGGTTTCTCTATAAATAGAGGCTGACATTCCTGTTCTGTTTCAAGATAT 7663
2y 7570 TCAATACCGCTCAGTATTTTAAATGAAGTCAATCTTAAGATTTGTTTGGGATCAATAGGA 7629
2b 7664 TCAATACCGCTCAGTATTTTAAATGAAGTCAATCTTAAGATTTGTTTGGGATCAATAGGA 7723
2y 7630 AAGCATATGAGCCCAACCAAGATGCAAAATGTTTGAATATGATGACCAAAATTTAAAGT 7689
2b 7724 AAGCATATGAGCCCAACCAAGATGCAAAATGTTTGAATATGATGACCAAAATTTAAAGT 7783
2y 7690 AGGAAAGTCAACCAACCACTCTCTTTCACTTAAAGTCTGCGCCGCAATCTGTAGGA 7749
2b 7784 AGGAAAGTCAACCAACCACTCTCTTTCACTTAAAGTCTGCGCCGCAATCTGTAGGA 7843
2y 7750 ACAAGCATGATCTTGTACTGTGATATTTTAAATATCCACAGTACT 7795
2b 7844 ACAAGCATGATCTTGTACTGTGATATTTTAAATATCCACAGTACT 7889

RESULT 4
US-10-447-161-8
; Sequence 8, Application US/10447161
; Publication No. US20040023314A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Rong-fu
; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis
; FILE REFERENCE: HO-P02484US1
; CURRENT APPLICATION NUMBER: US/10/447,161
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/383,530
; PRIOR FILING DATE: 2002-05-28
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8
; LENGTH: 8027
; TYPE: DNA
; ORGANISM: Human
US-10-447-161-8

Query Match 99.6%; Score 7760.2; DB 16; Length 8027;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 7779; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

2y 3 C T C G A G C C G G C T G T G T G C A C A G G G G A G A G G A A C C C C A G G C G C G A G C G G G A A G A 62
2b 3 G C C C G C C C G G C T G T G T G C A C A G G G G A G A G G A A C C C C A G G C G C G A G C G G G A A G 62
2y 63 G G G A C C T G C A G C C A C A C T T C T G T G C C T G C A T C C C T T C T G C C C T C C C C C C C C C 122
2b 63 G G G A C C T G C A G C C A C A C T T C T G T G C C T G C A T C C C T T C T G C C C T C C C C C C C C C 122
2y 123 C C T T C C C C A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C 182
2b 123 C C T T C C C C A C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C 182
2y 183 G G A T T T T T C C C G A G T T G C A G G G A G A C A A C T T G T G G G A C A C T T G C C C C G T G C 242
2b 183 G G A T T T T T C C C G A - G T T G C A A G G G A G A A A C T T G T G G C A A C T T G C C C C C G T G C 241
2y 243 G G G C G T C T C C C C A C C G T C T C A A C A T G C T T A G G G T C C G G G C C C G G G C T G C T G C 302
2b 242 G G G C G T C T C T C C C C A C C G T C T C A A C A T G C T T A G G G T C C G G G C C C G G G C T G C T G C 301

```

```

303 T G G C G T C C T G T G C C T G G G A C A G C G G T G C C C T C C A C G G G A G C C T C G A A G A G C A A G A G C C 362
302 T G G C G T C C A G T G C C T G G G A C A G C G G T G C C C T C C A C G G A G C C T C G A A G A G C A A G A G C C 361
363 A G G T C A C A A A T G G T T C A G C C C C A G T C C C G G T G G T G T C A G T C A A A A G A A G C C C C G G T T 422
362 A G G T C A C A A A T G G T T C A G C C C C A G T C C C G G T G G T G T C A G T C A A A A G A A G C C C C G G T T 421
423 G T T A T G A C A A T G G A A A C A C T A T C A G A T A A A T A A C A G T G G A G G G A G C C T A C C T A G G C A 482
422 G T T A T G A C A A T G G A A A C A C T A T C A G A T A A A T A A C A G T G G A G G C G A C C T A C C T A G G C A 481
483 A T G C C T T G G T T G T A C T T G T T A T G G A G A A C C C A G G T T T T A C T G C G A A A G T A A A C C T G 542
482 A T G C C T T G G T T G T A C T T G T T A T G G A G A A C C C A G G T T T T A C T G C G A G A G T A A A C C T G 541
543 A A G C T G A A G A C A C T T G C T T T G C A A A G T A C A C T G G G A C A C T T A C C G A G T G G T G C A C A C T T 602
542 A A G C T G A A G A C A C T T G C T T T G C A A A G T A C A C T G G G A C A C T T A C C G A G T G G T G C A C A C T T 601
603 A T G A G C G T C C T A A A G A C T C C A T G A T C T G G A C T G T A C C T G C A T C G G G C T G G G C G A G G G A 662
602 A T G A G C G T C C T A A A G A C T C C A T G A T C T G G A C T G T A C C T G C A T C G G G C T G G G C G A G G G A 661
663 G A A T A A G C T G T A C C A T C G C A A A C C G C T G C C A T G A A G G G G T C A G T C C T A C A A G A T T G T G 722
662 G A A T A A G C T G T A C C A T C G C A A A C C G C T G C C A T G A A G G G G T C A G T C C T A C A A G A T T G T G 721
723 A C A C C T G G A G A G A C C A C A T G A G A C T G G T G T T A C A T G T A G A G T G T G T G T C T T G G T A 782
722 A C A C C T G G A G A G A C C A C A T G A G A C T G G T G T T A C A T G T A G A G T G T G T G T G T G T G T A 781
783 A T G G A A A G G A A T G G A C C T G C A A G C C C A G C T A G C T A G A A G T T T T G A T C A T G C T G C T G 842
782 A T G G A A A G G A A T G G A C C C T G C A A G C C C A T A G C T A G A A G T T T T G A T C A T G C T G C T G 841
843 G G A C T T C C T A T G T G T G C G A G A A A C G T G G A G A A C C C T A C C A A G C C T G G A T G A T G T G T A G 902
842 G G A C T T C C T A T G T G T G C G A G A A A C G T G G A G A A C C C T A C C A G C C T G G A T G A T G T G T A G 901
903 A T T G A C T T G C C T G G G A A G C A C G G A C C G A C C A T C A C T T G C A C T T C T A G A A A T A G A T G C A 962
902 A T T G A C T T G C C T G G G A A G C A C G G A C C G A C C A T C A C T T G C A C T T C T A G A A A T A G A T G C A 961
963 A C G A T C A G A C A C A G G A C A C C T A T A G A A T T G G A G A C A C C T G G A G A G A G A G A T A A T C 1022
962 A C G A T C A G A C A C A G G A C A C C T A T A G A A T T G G A G A C A C C T G G A G A G A G A G A T A A T C 1021
1023 G A G G A A A C C T G C T C C A G T G C A T C T G C A G A G C A A C G G C C G A G A G A G T G G A A G T G T G A G A 1082
1022 G A G G A A A C C T G C T C C A G T G C A T C T G C A G A G C A A C G G C C G A G A G A T G G A A G T G T G A G A 1081
1083 G G C A C A C C T G T G C A G A C C A C A T C G A G G G A T C T G G C C C T T C A C C G A T G T T C C G T G C A G 1142
1082 G G C A C A C C T G T G C A G A C C A C A T C G A G G G A T C T G G C C C C T T C A C C G A T G T T C G T G C A G 1141
1143 C T G T T T A C C A A C C G A G C C T C A C C C C C A G C C T C C T C C C T A T G C C C A C T G T G T C A C A G A C A 1202
1142 C T G T T T A C C A A C C G A G C C T C A C C C C C A G C C T C C T C C C T A T G C C C A C T G T G T C A C A G A C A 1201
1203 G T G G T G T G T A C T C T G T G G G A T G C A G T G G C T G A A G A C A C A A G A A A T A A G C A A A T G C 1262
1202 G T G G T G T G T A C T C T G T G G G A T G C A G T G G C T G A A G A C A C A A G A A A T A A G C A A A T G C 1261
1263 T T T G C A C G T C C T G G C A A C G G A G T C A G T G C C A A G A G A C A G C T G T A A C C C A G A C T T A G C 1322
1262 T T T G C A C G T C C T G G C A A C G G A G T C A G T G C C A A G A G A C A G C T G T A A C C C A G A C T T A G C 1321
1323 G T G G C A A C T C A A A T G G A G A C C A T G T G T T T A C C A T T C A C T A A T G C A A T G C A G A C G T T C T 1382
1322 G T G G C A A C T C A A A T G G A G A C C A T G T G T C T T A C C A T T C A C C A T T A C C A A T G C A G A C G T T C T 1381

```

1383 ACTCTGACACAGAGGGGACAGGACGGACATCTTTGGTGCAGCACAACTTCGAATT 1442  
 1382 ACTCTGACACAGGAGGGGACAGGACGGACATCTTTGGTGCAGCACAACTTCGAATT 1441  
 1443 ATGACAGGACAGAAATACCTTTCTGACAGACCACTGTTTGGTTCAGACTCGAG 1502  
 1442 ATGACAGGACAGAAATACCTTTCTGACAGACCACTGTTTGGTTCAGACTCGAG 1501  
 1503 GAGGAAATCCAAATGGTGTGCTTGTGCTTCCCTTCTTATACAAACACAAATTACA 1562  
 1502 GAGGAAATCCAAATGGTGTGCTTGTGCTTCCCTTCTTATACAAACACAAATTACA 1561  
 1563 CTGATTGCACTTCTGAGGACAGAGCAACATGAGTGTGTTGGGACCAACAGAACT 1622  
 1562 CTGATTGCACTTCTGAGGACAGAGCAACATGAGTGTGTTGGGACCAACAGAACT 1621  
 1623 ATGATGCCACAGAAATTTGGGTTCTGCCCCATGGCTGCCACAGAGAAATCTGCACAA 1682  
 1622 ATGATGCCACAGAAATTTGGGTTCTGCCCCATGGCTGCCACAGAGAAATCTGCACAA 1681  
 1683 CCAATGAGGGTCACTGATCCGATTTGGGATGAGTGGGATTAAGCAGATGACATGGGTC 1742  
 1682 CCAATGAGGGTCACTGATCCGATTTGGGATGAGTGGGATTAAGCAGATGACATGGGTC 1741  
 1743 ACATGATGAGTGCACGTGTGTTGGGATGAGTGGGATTAAGCAGATGACATGGGTC 1802  
 1742 ACATGATGAGTGCACGTGTGTTGGGATGAGTGGGATTAAGCAGATGACATGGGTC 1801  
 1803 CGCAGTTTCAGATCAGTGCATTTGTTGATGACATCACTTACAAATGTTGAAACGACATTC 1862  
 1802 CGCAGTTTCAGATCAGTGCATTTGTTGATGACATCACTTACAAATGTTGAAACGACATTC 1861  
 1863 ACAGCGTTCATGAGAGGGGACATCTGCACTGTTGATGACATCACTTACAAATGTTGAAACGACATTC 1922  
 1862 ACAGCGTTCATGAGAGGGGACATCTGCACTGTTGATGACATCACTTACAAATGTTGAAACGACATTC 1921  
 1923 GGTGGAAGTGTGATCCGTCGACCAATGCGAGATTCAGAGCTGGGACGTTTATCAAA 1982  
 1922 GGTGGAAGTGTGATCCGTCGACCAATGCGAGATTCAGAGCTGGGACGTTTATCAAA 1981  
 1983 TTGGGATTTATGAGAGAGTATGTCATGTTGATGACATCACTTACAAATGTTGAAACGACATTC 2042  
 1982 TTGGGATTTATGAGAGAGTATGTCATGTTGATGACATCACTTACAAATGTTGAAACGACATTC 2041  
 2043 GTGSCATTTGGGAGTGGCATTTGCCAATTTTACAGACCTTATCCAAAGCTCAAGTGTCTG 2102  
 2042 GTGSCATTTGGGAGTGGCATTTGCCAATTTTACAGACCTTATCCAAAGCTCAAGTGTCTG 2101  
 2103 TCGAATTTATTTACATGAGACTCCGAGTCCGACCACTCCACCCCATCCAGTGGGATG 2162  
 2102 TCGAATTTATTTACATGAGACTCCGAGTCCGACCACTCCACCCCATCCAGTGGGATG 2161  
 2163 CACACAGCCATCTCACATTTCCAAATTTTACAGACCTTATCCAAAGCTCAAGTGTCTG 2222  
 2162 CACACAGCCATCTCACATTTCCAAATTTTACAGACCTTATCCAAAGCTCAAGTGTCTG 2221  
 2223 GCGTTTGGAGGAGTACATACAGGAGTCACTTAACTTCTACCACTCAAGGCTTGA 2282  
 2222 GCGTTTGGAGGAGTACATACAGGAGTCACTTAACTTCTACCACTCAAGGCTTGA 2281  
 2283 AGCTGTGTGTATAGAGGGCCAGTCTCATAGCATCCAGCATCCAGGATCCGACCAAGAG 2342  
 2282 AGCTGTGTGTATAGAGGGCCAGTCTCATAGCATCCAGCATCCAGGATCCGACCAAGAG 2341  
 2343 TGACTGCTTTGATTTACACACAGGAGTCACTTAACTTCTACCACTCAAGGCTTGA 2402  
 2342 TGACTGCTTTGATTTACACACAGGAGTCACTTAACTTCTACCACTCAAGGCTTGA 2401  
 2403 CAGGAGAGCAGTCTCCCTTTTCTCTTGTGGCCACTTCTGAATCTGTGACCGAAATCA 2462  
 2402 CAGGAGAGCAGTCTCCCTTTTCTCTTGTGGCCACTTCTGAATCTGTGACCGAAATCA 2461  
 2463 CAGCCAGTAGCTTTGTGGTCTCTCTGGTCTCCGACCTCCGACCTGTCGGGATTCGGG 2522

2462 CAGCCAGTAGCTTTGTGGTCTCTCTGGTCTCCGACACCTGTCGGGATTCGGG 2521  
 2523 TGGAAATATGAGCTGAGTGCAGGAGGAGATGAGCCACAGTACCTGATCTTCCAAAGCAG 2582  
 2522 TGGAAATATGAGCTGAGTGCAGGAGGAGATGAGCCACAGTACCTGATCTTCCAAAGCAG 2581  
 2583 CACTTCTGTGAAACATCCCTGACCTGCTTCTTCCGCGGAAATACATTTGTAATGCTATC 2642  
 2582 CCACTTCTGTGAAACATCCCTGACCTGCTTCTTCCGCGGAAATACATTTGTAATGCTATC 2641  
 2643 AGATATCTGAGGATGGGAGCAGAGTTTGTATCTGTCTACTTTCACAAAACACAGGCTG 2702  
 2642 AGATATCTGAGGATGGGAGCAGAGTTTGTATCTGTCTACTTTCACAAAACACAGGCTG 2701  
 2703 ATGCCCTCTGTGACCCGACTGTGGACCAAGTTGATGACACCTCAATTTGTTGCTGGA 2762  
 2702 ATGCCCTCTGTGACCCGACTGTGGACCAAGTTGATGACACCTCAATTTGTTGCTGGA 2761  
 2763 GCAGACCCAGGCTCCCATCACAGGATCAGAAATAGTCTTATCGCCATCAGTAGAGGTA 2822  
 2762 GCAGACCCAGGCTCCCATCACAGGATCAGAAATAGTCTTATCGCCATCAGTAGAGGTA 2821  
 2823 GCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGCTCACTCAGTACTTGCAC 2882  
 2822 GCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGCTCACTCAGTACTTGCAC 2881  
 2883 CTGGTGTCAAGTAAATCATCATCTATCTGCTGGAAGAAATCAAGAAAGTACACCTG 2942  
 2882 CTGGTGTCAAGTAAATCATCATCTATCTGCTGGAAGAAATCAAGAAAGTACACCTG 2941  
 2943 TTGTCAATTCACAAAGAAACCACTGGGACCCGCTCAGATACAGTGGCTCTCCAGGG 3002  
 2942 TTGTCAATTCACAAAGAAACCACTGGGACCCGCTCAGATACAGTGGCTCTCCAGGG 3001  
 3003 ACCTGCACTTTGTGGAAGTGCAGAGAGTGAAGTCACTATGTCGACACCCGCTGAGA 3062  
 3002 ACCTGCACTTTGTGGAAGTGCAGAGAGTGAAGTCACTATGTCGACACCCGCTGAGA 3061  
 3063 GTGCAGTGACCGGCTACCGTGTGATGATCCCTGCAACCTGCTGGGAGCAGCGGC 3122  
 3062 GTGCAGTGACCGGCTACCGTGTGATGATCCCTGCAACCTGCTGGGAGCAGCGGC 3121  
 3123 AGAGCTGCCCATCAGAGGAAACACTTTGACAGAGTCAACGCTGCTCCCTGGGCTCA 3182  
 3122 AGAGCTGCCCATCAGAGGAAACACTTTGACAGAGTCAACGCTGCTCCCTGGGCTCA 3181  
 3183 CCTATTACTTCAAAGTCTTTGCAAGTGCATCCGCTGAGAGGAGAGAGCCTCTGACTGCTC 3242  
 3182 CCTATTACTTCAAAGTCTTTGCAAGTGCATCCGCTGAGAGGAGAGAGCCTCTGACTGCTC 3241  
 3243 AACAGACAAACCAACTGGATGCTCCCACTAACTCCAGTTTGTCAATGAACCTGATCTTA 3302  
 3242 AACAGACAAACCAACTGGATGCTCCCACTAACTCCAGTTTGTCAATGAACCTGATCTTA 3301  
 3303 CTGTCTCTGAGATGAGACTCCACTCGGCGCCAGATAACAGGATACCGACTACCGGTG 3362  
 3302 CTGTCTCTGAGATGAGACTCCACTCGGCGCCAGATAACAGGATACCGACTACCGGTG 3361  
 3363 GCTTACCCGAGAGGACAGCCAGGAGTACATGTTGGGTCCCTCTGCTCCAGTACC 3422  
 3362 GCTTACCCGAGAGGAGCCAGGAGTACATGTTGGGTCCCTCTGCTCCAGTACC 3421  
 3423 CACTGAGGATCTGCAAGCTGCACTGAGTACACCGTATCCCTGCTGGCCATAAAGGGCA 3482  
 3422 CACTGAGGATCTGCAAGCTGCACTGAGTACACCGTATCCCTGCTGGCCATAAAGGGCA 3481  
 3483 ACCAAGAGAGCCCAAGGCACTGGAGTCTTTACCACTGTCAGCTGGAGTCTTATTC 3542  
 3482 ACCAAGAGAGCCCAAGGCACTGGAGTCTTTACCACTGTCAGCTGGAGTCTTATTC 3541  
 3543 CACTTACCAACCGAGGAGTGAAGTGAAGACCACTGATGATCAATGACGCTCTCCAA 3602

3542 CACCTTACACACCGAGGTGACTGAGACACCAATCGTGATCAATGGAGCGCTGCTCCAA 3601  
 3603 GAATTTGGTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGACCAACGAGAAAGTGA 3662  
 3602 GAATTTGGTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGACCAACGAGAAAGTGA 3661  
 3663 CTTTCAGACTCAGGAAGCATCGTTGTGTCCGGCTTGACTCCAGAGTAGAATACGCTTACA 3722  
 3662 CTTTCAGACTCAGGAAGCATCGTTGTGTCCGGCTTGACTCCAGAGTAGAATACGCTTACA 3721  
 3723 CCATCCAAAGTCTTGAGAGATGGACAGAGAAAGAGATGCGCAATTTGTAACAAAGTGGTGA 3782  
 3722 CCATCCAAAGTCTTGAGAGATGGACAGAGAAAGAGATGCGCAATTTGTAACAAAGTGGTGA 3781  
 3783 CACCATTTGCTCCACCAACAACTTGATCTGGAGGCAACCCCTGACCTGGAGTGGCTCA 3842  
 3782 CACCATTTGCTCCACCAACAACTTGATCTGGAGGCAACCCCTGACCTGGAGTGGCTCA 3841  
 3843 CAGTCTCTCTGGAGAGGACCAACCCAGACATTTACTGGTTATAGAAATTTACCACAACCC 3902  
 3842 CAGTCTCTCTGGAGAGGACCAACCCAGACATTTACTGGTTATAGAAATTTACCACAACCC 3901  
 3903 CTAACAACGGCAGCAGGAAATTTCTTTGGAAGAGTGGTCCATGCTGATCAGAGTCCCT 3962  
 3902 CTAACAACGGCAGCAGGAAATTTCTTTGGAAGAGTGGTCCATGCTGATCAGAGTCCCT 3961  
 3963 GCATTTTGATTAACCTGAGTCCCGGCTGGAGTACAAATGTCACTGTTTACACTGTCAAGG 4022  
 3962 GCATTTTGATTAACCTGAGTCCCGGCTGGAGTACAAATGTCACTGTTTACACTGTCAAGG 4021  
 4023 ATGACAGGAAAGTGTCCCTATCTCTGATACATCATCCAGCTGTTCCTCCCACTG 4082  
 4022 ATGACAGGAAAGTGTCCCTATCTCTGATACATCATCCAGCTGTTCCTCCCACTG 4081  
 4083 ACCTGCGATTAACCAACATTTGCTCCAGACACCATGCTGTCACCTGGGCTCCACCCCAT 4142  
 4082 ACCTGCGATTAACCAACATTTGCTCCAGACACCATGCTGTCACCTGGGCTCCACCCCAT 4141  
 4143 CCATTTGATTAACCAACTCTCTGGTGGTTACTCACTGTGAAATAATGAGAAAGATGTTG 4202  
 4142 CCATTTGATTAACCAACTCTCTGGTGGTTACTCACTGTGAAATAATGAGAAAGATGTTG 4201  
 4203 CAGAGTTGTCATTTCTCTCCATCAGCAATGAGTGTCCTTAAACAAATCTCTGCTGGTA 4262  
 4202 CAGAGTTGTCATTTCTCTCCATCAGCAATGAGTGTCCTTAAACAAATCTCTGCTGGTA 4261  
 4263 CAGAAATATGATGATGTCTCCAGTGTCTACGAAACCAATGAGAGCAACCTCTTAGAG 4322  
 4262 CAGAAATATGATGATGTCTCCAGTGTCTACGAAACCAATGAGAGCAACCTCTTAGAG 4321  
 4323 GAAGACAGAAACAGGCTTGATTTCCCACTGGCATTTCTTGATATTAAGTCCCA 4382  
 4322 GAAGACAGAAACAGGCTTGATTTCCCACTGGCATTTCTTGATATTAAGTCCCA 4381  
 4383 ACTCTTTTACTGTGCACTGATTTCTCTCGAGCACCACATCACTGGCTACAGGATCGCC 4442  
 4382 ACTCTTTTACTGTGCACTGATTTCTCTCGAGCACCACATCACTGGCTACAGGATCGCC 4441  
 4443 ATCATCCCGAGCATTTAGTGGGAGACTCGAGAAGATCGGGTCCCACTCTCGGAATT 4502  
 4442 ATCATCCCGAGCATTTAGTGGGAGACTCGAGAAGATCGGGTCCCACTCTCGGAATT 4501  
 4503 CCATCACCTCACCACCTCACTCCAGGCAAGATATGGTCAGCATCGTTGCTCTTA 4562  
 4502 CCATCACCTCACCACCTCACTCCAGGCAAGATATGGTCAGCATCGTTGCTCTTA 4561  
 4563 ATGGCAGAGAGAAAGTCCCTTTATGATGGCCAAACAAATCAAGTTTCTGATGTTCCGA 4622  
 4562 ATGGCAGAGAGAAAGTCCCTTTATGATGGCCAAACAAATCAAGTTTCTGATGTTCCGA 4621  
 4623 GGGACCTGGAAGTTGTTGCTGGGACCCCAACAGCCTACTGATCAGCTGGAGTCTCTG 4682  
 4622 GGGACCTGGAAGTTGTTGCTGGGACCCCAACAGCCTACTGATCAGCTGGAGTCTCTG 4681

QY 4683 CTGTCAAGTGTGATATTTACAGGATCACTTACGGGAAACAGGAGAAATAGCCCTGTC 4742  
 DB 4682 CTGTCAAGTGTGATATTTACAGGATCACTTACGGGAAACAGGAGAAATAGCCCTGTC 4741  
 QY 4743 AGGAGTTTCACTGTGCTGGGAGCAAGTCTACAGTACCATCAGCGGCTTTAAACCTGGAG 4802  
 DB 4742 AGGAGTTTCACTGTGCTGGGAGCAAGTCTACAGTACCATCAGCGGCTTTAAACCTGGAG 4801  
 QY 4803 TTGATTTATACCATCACTGTGTATCTGTCACTGCGCTGGAGACAGCCCGCAGAGAGCA 4862  
 DB 4802 TTGATTTATACCATCACTGTGTATCTGTCACTGCGCTGGAGACAGCCCGCAGAGAGCA 4861  
 QY 4863 AGCCAAATTTCCATTAATTTACCCGAATTTGACAAATTTGACAAATTTCCAGATGCAAGTGACCG 4922  
 DB 4862 AGCCAAATTTCCATTAATTTACCCGAATTTGACAAATTTGACAAATTTCCAGATGCAAGTGACCG 4921  
 QY 4923 ATGTTTCAGGACAACAGCATTTAGTGTCAAGTGGCTGCTTCAAGTTCCCTGTTACTGGTT 4982  
 DB 4922 ATGTTTCAGGACAACAGCATTTAGTGTCAAGTGGCTGCTTCAAGTTCCCTGTTACTGGTT 4981  
 QY 4983 ACAGAGTTAAACCAACCACTCCCAAAATGGACACAGGACCAACAAACAACTAAACCTGCAGTGC 5042  
 DB 4982 ACAGAGTTAAACCAACCACTCCCAAAATGGACACAGGACCAACAAACAACTAAACCTGCAGTGC 5041  
 QY 5043 CAGATCAAAACAGAAATGACTATTTGAAGCTTTGAGCTCCACAGTGGAGTATGTGTAGTG 5102  
 DB 5042 CAGATCAAAACAGAAATGACTATTTGAAGCTTTGAGCTCCACAGTGGAGTATGTGTAGTG 5101  
 QY 5103 TCTATGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGTTTCAAGTGCAGTACCAACAA 5162  
 DB 5102 TCTATGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGTTTCAAGTGCAGTACCAACAA 5161  
 QY 5163 TTGATTCGCCCTAAAGGACTGCAATTTCACTGATGTGATGTGATTCATCAAAATTTGCTT 5222  
 DB 5162 TTGATTCGCCCTAAAGGACTGCAATTTCACTGATGTGATGTGATTCATCAAAATTTGCTT 5221  
 QY 5223 GGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGTCAGCTCTCTCGAGCCTCGAGGATG 5282  
 DB 5222 GGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGTCAGCTCTCTCGAGCCTCGAGGATG 5281  
 QY 5283 GAATCCATGAGCTATTTCCCTGCACCTGATGTGTGAGAAAGACACTGCAGAGTGCAGAGGCC 5342  
 DB 5282 GAATCCATGAGCTATTTCCCTGCACCTGATGTGTGAGAAAGACACTGCAGAGTGCAGAGGCC 5341  
 QY 5343 TCAGACCGGTTTCTGAGTACACAGTCACTGTGTGTGCTTGCAGCATGATATGGAGGCC 5402  
 DB 5342 TCAGACCGGTTTCTGAGTACACAGTCACTGTGTGTGCTTGCAGCATGATATGGAGGCC 5401  
 QY 5403 AGCCCTGATTTGGAACCCAGTCCACAGCTATTTCTGCAACCAACTGACCTGAAATTTGCTC 5462  
 DB 5402 AGCCCTGATTTGGAACCCAGTCCACAGCTATTTCTGCAACCAACTGACCTGAAATTTGCTC 5461  
 QY 5463 AGGTCAACCCCAAGGCTGAGCGCCAGTGGACACCAACCCCAATGTTTCAAGTCTAGT 5522  
 DB 5462 AGGTCAACCCCAAGGCTGAGCGCCAGTGGACACCAACCCCAATGTTTCAAGTCTAGT 5521  
 QY 5523 ATCGAGTGGGTTGACCCCAAGGAGAGACCGGACCAATGAAAGAAATCAACCTTTGCTC 5582  
 DB 5522 ATCGAGTGGGTTGACCCCAAGGAGAGACCGGACCAATGAAAGAAATCAACCTTTGCTC 5581  
 QY 5583 CTGACAGCTCATCCGTTGTTGATCAGGACTTATGGTGGCCACCAAAATATGAAGTGTG 5642  
 DB 5582 CTGACAGCTCATCCGTTGTTGATCAGGACTTATGGTGGCCACCAAAATATGAAGTGTG 5641  
 QY 5643 TCTATGCTCTTAAGGACACTTTTGAACAGGAGAGACCGGACCAATGAAAGAAATCAACCTTTGCTC 5702  
 DB 5642 TCTATGCTCTTAAGGACACTTTTGAACAGGAGAGACCGGACCAATGAAAGAAATCAACCTTTGCTC 5701  
 QY 5703 AGAATGTACGCCCAACCAAGAGAGGCTCGTGTGACAGATGCTACTGAGACCAACCATCA 5762  
 DB 5702 AGAATGTACGCCCAACCAAGAGAGGCTCGTGTGACAGATGCTACTGAGACCAACCATCA 5761



5763	TTAGCTGGAGAACCAAGA	CTGAGACGATCACTGGCCT	CCAAAGTTGAATGCGGTTCCAGCCA	5822	
5762	TTAGCTGGAGAACCAAGA	CTGAGACGATCACTGGCTT	CCAAAGTTGAATGCGGTTCCAGCCA	5821	
5823	ATGGCCAGACTCCAAT	TCGAGAGAACCATCAAGCCAGATGTCAGAGCTACACCATCACATG	5882		
5822	ATGGCCAGACTCCAAT	TCGAGAGAACCATCAAGCCAGATGTCAGAGCTACACCATCACAG	5881		
5883	GCTTCAACACGAGCACTGACTA	CAAGATCTACCTGTACACCTTTGAA	TGACAAATGCTCGGA	5942	
5882	GTTTCAACACGAGCACTGACTA	CAAGATCTACCTGTACACCTTTGAA	TGACAAATGCTCGGA	5941	
5943	GCTCCCTGTGTCATCGAGCCCTCCACT	TGCCATTCATGCACCATCCACCTCGGTTTCC	6002		
5942	GCTCCCTGTGTCATCGAGCCCTCCACT	TGCCATTCATGCACCATCCACCTCGGTTTCC	6001		
6003	TGGCCACCACACCAATTCCTT	CGCTGGTATCATGGCAGCGCCACGCTGGCAGGATTACCG	6062		
6002	TGGCCACCACACCAATTCCTT	CGCTGGTATCATGGCAGCGCCACGCTGGCAGGATTACCG	6061		
6063	GCTACATCATCAAGTATGAGAGCCCTGGTCTCT	CTCCAGAGAAAGTGTTCCCTCGCGCCC	6122		
6062	GCTACATCATCAAGTATGAGAGCCCTGGTCTCT	CTCCAGAGAAAGTGTTCCCTCGCGCCC	6121		
6123	GCCTCGTGTCCAGAGGCTACTATTACT	TGGCTGGAA	CCGGAA	CGGAATATACAATTT	6182
6122	GCCTCGTGTCCAGAGGCTACTATTACT	TGGCTGGAA	CCGGAA	CGGAATATACAATTT	6181
6183	ATGTCAATTGCCCTGGAAGATPAAT	TCAGAGAGCGAGCCCTGATTTGGAAGAAAGACAG	6242		
6182	ATGTCAATTGCCCTGGAAGATPAAT	TCAGAGAGCGAGCCCTGATTTGGAAGAAAGACAG	6241		
6243	ACGAGCTTCCCCAACTGGTAACCTTCCAC	CCCCCAATCTTCATGACACGAGATCTTGG	6302		
6242	ACGAGCTTCCCCAACTGGTAACCTTCCAC	CCCCCAATCTTCATGACACGAGATCTTGG	6301		
6303	ATGTTCTCTTCCACAGTTCAAAAGACCCCTTTC	GTCA	CCCCACCTGGGTATGACACTGGAA	6362	
6302	ATGTTCTCTTCCACAGTTCAAAAGACCCCTTTC	GTCA	CCCCACCTGGGTATGACACTGGAA	6361	
6363	ATGGTATTGAGTTCCTGGCACTTCTGGTCAGCAA	CCCCAGTGTGGGCAACAATGATCT	6422		
6362	ATGGTATTGAGTTCCTGGCACTTCTGGTCAGCAA	CCCCAGTGTGGGCAACAATGATCT	6421		
6423	TTGAGGAAATGTTTTAGCGGACCA	CCGCCCCACACCGCCCATTAAGCATA	6482		
6422	TTGAGGAAATGTTTTAGCGGACCA	CCGCCCCACACCGCCCATTAAGCATA	6481		
6483	GGCCAGA	CCATACCGCCGAATGTAGGACAAAGAGCTCTCTCTCAGACCAACCATCTCAT	6542		
6482	GGCCAGA	CCATACCGCCGAATGTAGGACAAAGAGCTCTCTCTCAGACCAACCATCTCAT	6541		
6543	GGGCCCATTCACAGGACACTTCTGAGTACAT	CATTTCATCTCTGTGGCACTGATG	6602		
6542	GGGCCCATTCACAGGACACTTCTGAGTACAT	CATTTCATCTCTGTGGCACTGATG	6601		
6603	AGAA	CCCTTACAGTTACAGGTTCTCGAACTCTTACCAAGTCCACTCTGACAGGCTCA	6662		
6602	AGAA	CCCTTACAGTTACAGGTTCTCGAACTCTTACCAAGTCCACTCTGACAGGCTCA	6661		
6663	CCAGAGTGCCACTACAACATCATAGT	GAGGCACTGAAAGACAGCAGAGCATTAAGG	6722		
6662	CCAGAGTGCCACTACAACATCATAGT	GAGGCACTGAAAGACAGCAGAGCATTAAGG	6721		
6723	TTCCGGGAAGAGTTGTTA	CCGTGGGCAACTCTGTCAACGAAGGCTTGAA	CCAACTTACGG	6782	
6722	TTCCGGGAAGAGTTGTTA	CCGTGGGCAACTCTGTCAACGAAGGCTTGAA	CCAACTTACGG	6781	
6783	ATGACTCGTCTTTGACCCCTTACACAGTG	TGCCATATGCCGTGGAGATGAGTGGGAAC	6842		
6782	ATGACTCGTCTTTGACCCCTTACACAGTG	TGCCATATGCCGTGGAGATGAGTGGGAAC	6841		
6843	GAATGTCTGAATCAGGCTTTTAAACTGT	TGTGCCAGTGTCTTAGGCTTTGGAAAGTGGTCAAT	6902		

Db	6842	GAATGCTGTAATCAGGCTTTAAACTCTTTGTGCCAGTGCCTTAGGCTTTGGAGTGTCACTT	6901
Qy	6903	TCAGATGTGATTCATCTAGATGCTGCATGACAATGGTGTGAACCTACAAGATTGGAGAGA	6962
Db	6902	TCAGATGTGATTCATCTAGATGCTGCATGACAATGGTGTGAACCTACAAGATTGGAGAGA	6961
Qy	6963	AGTGGGACCGTCAGGAGGAAATGGCCAGATGATGAGCTGCACATGCTCTTGGGAAACGGAA	7022
Db	6962	AGTGGGACCGTCAGGAGGAAATGGCCAGATGATGAGCTGCACATGCTCTTGGGAAACGGAA	7021
Qy	7023	AAGGAGAAATCAAGTGTGACCCCTCATGAGGCAACGTGTTACGATGATGGGAACATACC	7082
Db	7022	AAGGAGAAATCAAGTGTGACCCCTCATGAGGCAACGTGTTACGATGATGGGAACATACC	7081
Qy	7083	ACGTAGGAGAACAGTGGCAGAGGAATATCTCGTGCCTATTTGCTCCTGCACATGCTTTG	7142
Db	7082	ACGTAGGAGAACAGTGGCAGAGGAATATCTCGTGCCTATTTGCTCCTGCACATGCTTTG	7141
Qy	7143	GAGGCCACGCGGGCTGGCGCTGTGACAACTGCGCAGACCTGGGGGTGAACCCAGTCCCG	7202
Db	7142	GAGGCCACGCGGGCTGGCGCTGTGACAACTGCGCAGACCTGGGGGTGAACCCAGTCCCG	7201
Qy	7203	AAGGCACTACTGGCCAGTCTCTACAAACAGTATTCTCAGAGATACCATCAGAGAAACAAACA	7262
Db	7202	AAGGCACTACTGGCCAGTCTCTACAAACAGTATTCTCAGAGATACCATCAGAGAAACAAACA	7261
Qy	7263	CTAATGTTAAATGCCAAATGAGTGTCTCATGCGCTTTAGATGTPACAGGTGACAGAGAAG	7322
Db	7262	CTAATGTTAAATGCCAAATGAGTGTCTCATGCGCTTTAGATGTPACAGGTGACAGAGAAG	7321
Qy	7323	ATTCCCGAGAGTAAATCATCTTTCCAAATCCAGAGGAAACAGCATGTCTCTCTCCCAAGAT	7382
Db	7322	ATTCCCGAGAGTAAATCATCTTTCCAAATCCAGAGGAAACAGCATGTCTCTCTCCCAAGAT	7381
Qy	7383	CCATCTAAACTGGAGTGATTTAGCAGACCCAGCTTTAGAGTTCTTTCTTTCTTTTAAGC	7442
Db	7382	CCATCTAAACTGGAGTGATTTAGCAGACCCAGCTTTAGAGTTCTTTCTTTCTTTTAAGC	7441
Qy	7443	CCTTTGCTCTGGAGGAGTTCTCCAGCTTCAGCTCAACTCACAAGCTCTCCCAAGCATCAC	7502
Db	7442	CCTTTGCTCTGGAGGAGTTCTCCAGCTTCAGCTCAACTCACAAGCTCTCCCAAGCATCAC	7501
Qy	7503	CCTGGGAGTTTCCCTGAGGGTTTTCTCATAAATGAGGGCTGCACATTCGCTGTCTTC	7562
Db	7502	CCTGGGAGTTTCCCTGAGGGTTTTCTCATAAATGAGGGCTGCACATTCGCTGTCTTC	7561
Qy	7563	GNAGTATTCAATACCGCTCAGTATTTTAATGAAGTGATTTCTAAGATTTGGTTTGGGATC	7622
Db	7562	GNAGTATTCAATACCGCTCAGTATTTTAATGAAGTGATTTCTAAGATTTGGTTTGGGATC	7621
Qy	7623	AATAGGAAAGCATATGCAGGCCAACCAAGATGCAAAATGTTTTGAAATGATATGCCAAAAT	7682
Db	7622	AATAGGAAAGCATATGCAGGCCAACCAAGATGCAAAATGTTTTGAAATGATATGCCAAAAT	7681
Qy	7683	TTTTAAGTAGGAAGATCACCCAAACATCTCTGCTTTTCACTTTAAGTGTCTGCCCGCATAC	7742
Db	7682	TTTTAAGTAGGAAGATCACCCAAACATCTCTGCTTTTCACTTTAAGTGTCTGCCCGCATAC	7741
Qy	7743	TGTTAGGAACAGCATGATCTTTGTTACTGTGATATTTTAAATATCCACAGTACT	7795
Db	7742	TGTTAGGAACAGCATGATCTTTGTTACTGTGATATTTTAAATATCCACAGTACT	7794

```

RESULT 5
US-10-098-841-7
; Sequence 7, Application US/10098841
; Publication NO. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chonghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun

```



APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yungling  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhao, Qing A.  
APPLICANT: Ren, Feiyan  
APPLICANT: Chen, Rui-hong  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: Wehrman, Tom  
APPLICANT: Zhang, Jie  
APPLICANT: Qian, Xiaohong B.  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 784CIP2  
CURRENT APPLICATION NUMBER: US/10/098,841  
CURRENT FILING DATE: 2002-03-13  
PRIOR APPLICATION NUMBER: 09/598,042  
PRIOR FILING DATE: 2000-06-20  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 331  
SOFTWARE: pt\_FL\_genes Version 1.0  
SEQ ID NO 7  
LENGTH: 8230  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (363)..(7523)  
IS-10-098-841-7

Query Match 98.4%; Score 7671.8; DB 13; Length 8230;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 7779; Conservative 0; Mismatches 7; Indels 93; Gaps 1;  
2y 10 CCGGCTGTGTGTCACAGGGGAGAGAGGAAACCCAGGCGCGAGCGGGAAGAGGGGACC 69  
2b 104 CCGGCTGTGTGTCACAGGGGAGAGAGGAAACCCAGGCGCGAGCGGGAAGAGGGGACC 163  
2y 70 TGGAGCACAACTTCTGTGCTCTGCATCCCTTCTGCTCCCTCCACCCGCTCCCTCC 129  
2b 164 TGCAGCACAACTTCTGTGCTCTGCATCCCTTCTGCTCCCTCCACCCGCTCCCTCC 223  
2y 130 CACCCTCTGGCCCCACCTTCTTGGAGGCGACAAACCCCGGGGCGCATTAAGAAGGATTT 189  
2b 224 CACCCTCTGGCCCCACCTTCTTGGAGGCGACAAACCCCGGGGCGCATTAAGAAGGATTT 283  
2y 190 TTCCCGCAGGTTGCGAGGGAAGCAAACTTGGTGGCAACTTGCCTCCCGGTGCGGGGTC 249  
2b 284 TTCCCGCAGGTTGCGAGGGAAGCAAACTTGGTGGCAACTTGCCTCCCGGTGCGGGGTC 343  
2y 250 TCTCCGCCACCGTCTCAACATGCTTAGGGGTCCGGGCGCGGGGCTGCTGTCTGGCGCT 309  
2b 344 TCTCCGCCACCGTCTCAACATGCTTAGGGGTCCGGGCGCGGGGCTGCTGTCTGGCGCT 403  
2y 310 CCTGTGCTGGGACAGCGGTGCCCTCCACGGGAGCCTTCGAAGAGCAAGAGGCGAGGCTCA 369  
2b 404 CCTGTGCTGGGACAGCGGTGCCCTCCACGGGAGCCTTCGAAGAGCAAGAGGCGAGGCTCA 463  
2y 370 GCAATGTTTCAGCCCCAGTCCCGGTGGCTGTGAGTCAGTCAAGCAAGCCCGGTGTTATGA 429  
2b 464 GCAATGTTTCAGCCCCAGTCCCGGTGGCTGTGAGTCAGTCAAGCAAGCCCGGTGTTATGA 523  
2y 430 CAATGGAAAACTATCAGATAAATCAACAGTGGGAGCGGACCTTACCTAGGCAATGGCTT 489  
2b 524 CAATGGAAAACTATCAGATAAATCAACAGTGGGAGCGGACCTTACCTAGGCAATGGCTT 583  
2y 490 GGTGTTGTAAGTATGAGGAGCCGAGGTTTAACTGCGAAAGTAAACCTGAGGCTGA 549  
2b 584 GGTGTTGTAAGTATGAGGAGCCGAGGTTTAACTGCGAGAGTAAACCTGAGGCTGA 643

QY 550 AGAGACTTGTCTTTGACAAGTACACTGGGACACTTACCGAGTGGGTGACACTTATCAGCG 609  
DB 644 AGAGACTTGTCTTTGACAAGTACACTGGGAACTTTACCGAGTGGGTGACACTTATGAGCG 703  
QY 610 TCCTAAAGACTTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGGCGAGGAGATAAG 669  
DB 704 TCCTAAAGACTTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGGCGAGGAGATAAG 763  
QY 670 CTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGTCCCTACAAGATTGGTGCACACTG 729  
DB 764 CTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGTCCCTACAAGATTGGTGCACACTG 823  
QY 730 GAGGAGACCAATGAGACTGTGTGTGTACATGTTAGAGTGTGTGTGTGTGTGTGTGTGT 789  
DB 824 GAGGAGACCAATGAGACTGTGTGTGTACATGTTAGAGTGTGTGTGTGTGTGTGTGTGT 883  
QY 790 AGGAGATGGACCTGCAAGCCCATAGCTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 849  
DB 884 AGGAGATGGACCTGCAAGCCCATAGCTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 943  
QY 850 CTATGTGTGTGCGAAGAACGCTGGGAGAAAGCCCTACCAAGGCTGGATGATGTTAGATTGATC 909  
DB 944 CTATGTGTGTGCGAAGAACGCTGGGAGAAAGCCCTACCAAGGCTGGATGATGTTAGATTGATC 1003  
QY 910 TTGCTGGGAGAGGCGAGCGGACCCATCTTGCATCTTAGAATATAGATGCAACGATCA 969  
DB 1004 TTGCTGGGAGAGGCGAGCGGACCCATCTTGCATCTTAGAATATAGATGCAACGATCA 1063  
QY 970 GGACACAGGACATCCCTATAGAAATGGAGACACCTGGAGCAAGAGGATATATCAGGAAA 1029  
DB 1064 GGACACAGGACATCCCTATAGAAATGGAGACACCTGGAGCAAGAGGATATATCAGGAAA 1123  
QY 1030 CTTGCTCCAGTGCATCTGCA CAGCAACCGCGGAGGAGAGTGGAAAGTGTGAGAGGCAAC 1089  
DB 1124 CTTGCTCCAGTGCATCTGCA CAGCAACCGCGGAGGAGAGTGGAAAGTGTGAGAGGCAAC 1183  
QY 1090 CTCTGTGAGACCCATCTGAGCGGATCTGGCCCTTCCACCGATGTTCGTGAGCTGTTTA 1149  
DB 1184 CTCTGTGAGACCCATCTGAGCGGATCTGGCCCTTCCACCGATGTTCGTGAGCTGTTTA 1243  
QY 1150 CCAACCGGACCTCACCCCCAGCCCTCTCCCTATGGCCACTGTGTACAGACAGTGTGT 1209  
DB 1244 CCAACCGGACCTCACCCCCAGCCCTCTCCCTATGGCCACTGTGTACAGACAGTGTGT 1303  
QY 1210 GGTCTACTCTGTGGGATGCAAGTGGCTGAAGACACAGGAAATAGCAAAATGCTTTGAC 1269  
DB 1304 GGTCTACTCTGTGGGATGCAAGTGGCTGAAGACACAGGAAATAGCAAAATGCTTTGAC 1363  
QY 1270 GTGCCCTGGCAACGAGTCAAGTGGCTGAAGACACAGGAAATAGCAAAATGCTTTGAC 1329  
DB 1364 GTGCCCTGGCAACGAGTCAAGTGGCTGAAGACACAGGAAATAGCAAAATGCTTTGAC 1423  
QY 1330 CTCAAATGGAGAGCAATGTGTCTTACCATTCACCTAGATGGCAGGAGTCTACTCTG 1389  
DB 1424 CTCAAATGGAGAGCAATGTGTCTTACCATTCACCTAGATGGCAGGAGTCTACTCTG 1483  
QY 1390 CACCAAGAGGCGGACAGGACCGGACATCTTTGGTGCAGACAACTTCGAATATGAGCA 1449  
DB 1484 CACCAAGAGGCGGACAGGACCGGACATCTTTGGTGCAGACAACTTCGAATATGAGCA 1543  
QY 1450 GGAACCAAGAAATACTCTTCTGTCAGACCA CACTGTTTTGTTTGTGAGCTCGAGGAGAAA 1509  
DB 1544 GGAACCAAGAAATACTCTTCTGTCAGACCA CACTGTTTTGTTTGTGAGCTCGAGGAGAAA 1603  
QY 1510 TTCCCATGTGCTTGTGCACTTCCCTTCCCTATACCAACCAACCAATATCACTGATTG 1569  
DB 1604 TTCCCATGTGCTTGTGCACTTCCCTTCCCTATACCAACCAACCAATATCACTGATTG 1663  
QY 1570 CACTTCTGAGGCGAGAGAGCAAA CATAAGTGTGTGTGGACCA CACAGAACTATGATGC 1629  
DB 1664 CACTTCTGAGGCGAGAGAGCAAA CATAAGTGTGTGTGGACCA CACAGAACTATGATGC 1723



Db 3884 GTCTCCACCAAACTTGCATCTGGAGGCAAACTCTGACACCTGGAGTGTCTACAGTCTC 3943  
 Qy 3850 CTGGGAGGAGGACCAACCCAGACATTAATCTGGTTATAGAAATACCAACACCCCTACAAA 3909  
 Db 3944 CTGGGAGGAGGACCAACCCAGACATTAATCTGGTTATAGAAATACCAACACCCCTACAAA 4003  
 Qy 3910 CGGCGACAGGGAATTTCTTTGGAAGAGTGGTCCATCTGCTGATCAGAGCTCTCTGCACCTT 3969  
 Db 4004 CGGCGACAGGGAATTTCTTTGGAAGAGTGGTCCATCTGCTGATCAGAGCTCTCTGCACCTT 4063  
 Qy 3970 TGATACTGAGTCCGCGCTGGAGTACAAATGTCAGTGTGTATGACTGTCAAGATGACAA 4029  
 Db 4064 TGATAACCTGATCCGCGCTGGAGTACAAATGTCAGTGTGTATGACTGTCAAGATGACAA 4123  
 Qy 4030 GGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGTCTCTCTCCCACTGACCTGCG 4089  
 Db 4124 GGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGTCTCTCTCCCACTGACCTGCG 4183  
 Qy 4090 ATTCAACCAATTTGGTCCAGACACCATGCTGTGTCACTGGGCTCCAGCCCATCATTTGA 4149  
 Db 4184 ATTCAACCAATTTGGTCCAGACACCATGCTGTGTCACTGGGCTCCAGCCCATCATTTGA 4243  
 Qy 4150 TTTAAACCAATTTCTGGTGGCTTACTCACTCTGTGAATAATGAGGAAGATGTTGAGAGTT 4209  
 Db 4244 TTTAAACCAATTTCTGGTGGCTTACTCACTCTGTGAATAATGAGGAAGATGTTGAGAGTT 4303  
 Qy 4210 GTCAATTTCTCTTCAGACAAATGAGTGGTCTTAACAAATCTCTGCTGTGTACAGAAAT 4269  
 Db 4304 GTCAATTTCTCTTCAGACAAATGAGTGGTCTTAACAAATCTCTGCTGTGTACAGAAAT 4363  
 Qy 4270 TGTAGTGTCTCTCCAGTGTCTAGACACATGAGAGACACCTCTTAGAGGAAGACA 4329  
 Db 4364 TGTAGTGTGTCTCTCCAGTGTCTAGACACATGAGAGACACCTCTTAGAGGAAGACA 4423  
 Qy 4330 GAAAAACAGTCTTGAATTTCCCACTGGCAATGACTTTTCTGATATTAATGCAACTCTTT 4389  
 Db 4424 GAAAAACAGTCTTGAATTTCCCACTGGCAATGACTTTTCTGATATTAATGCAACTCTTT 4483  
 Qy 4390 TACTGTGACGTGATGCTCTCGAGCCACCATCACTGGCTACAGGATCGGCATCATCC 4449  
 Db 4484 TACTGTGACGTGATGCTCTCGAGCCACCATCACTGGCTACAGGATCGGCATCATCC 4543  
 Qy 4450 CGAGCACTTCAGTGGGAGACCTCTCAGAAAGATCGGGTGCCCACTCTCGGAATTCATCAC 4509  
 Db 4544 CGAGCACTTCAGTGGGAGACCTCTCAGAAAGATCGGGTGCCCACTCTCGGAATTCATCAC 4603  
 Qy 4510 CCTCAACCACTCACTCAGGACACAGATATGTTGCTGAGCATGTTGCTCTTAATGGCAG 4569  
 Db 4604 CCTCAACCACTCACTCAGGACACAGATATGTTGCTGAGCATGTTGCTCTTAATGGCAG 4663  
 Qy 4570 AGAGGAAAGTCCCTTATTTGATTTGGCAACAAATCAACAGTTTCTGATGTTCCGAGGGACCT 4629  
 Db 4664 AGAGGAAAGTCCCTTATTTGATTTGGCAACAAATCAACAGTTTCTGATGTTCCGAGGGACCT 4723  
 Qy 4630 GGAAGTGTGCTGGGACCCCAACAGCTCTACTGATGATGCTGGGATGCTCTGCTGTAC 4689  
 Db 4724 GGAAGTGTGCTGGGACCCCAACAGCTCTACTGATGATGCTGGGATGCTCTGCTGTAC 4783  
 Qy 4690 AGTGAGATATTACAGGATCACTTACGGAGAAACAGAGGAAATAGCCCTGTCCAGGAGTT 4749  
 Db 4784 AGTGAGATATTACAGGATCACTTACGGAGAAACAGAGGAAATAGCCCTGTCCAGGAGTT 4843  
 Qy 4750 CACTGTGCTGGGACCAAGTCTTACAGCTACCATCAGCGGCTTTAAACCTGGAGTTGATTA 4809  
 Db 4844 CACTGTGCTGGGACCAAGTCTTACAGCTACCATCAGCGGCTTTAAACCTGGAGTTGATTA 4903  
 Qy 4810 TACCATCACTGTGATGCTGATCTGCTGGCGGTGGAGACAGCCCGCAAGCAGCAAGCCAAAT 4869  
 Db 4904 TACCATCACTGTGATGCTGATCTGCTGGCGGTGGAGACAGCCCGCAAGCAGCAAGCCAAAT 4963  
 Qy 4870 TTCCATTAATACCGAAACAGAAATTTGACAAACCAATCCAGATGCAAGTGACCGATGTTCA 4929  
 Db 4964 TTCCATTAATACCGAAACAGAAATTTGACAAACCAATCCAGATGCAAGTGACCGATGTTCA 5023

Qy 4930 GGACAAACAGCATTAGTGTCAAGTGGCTGCTTCAAGTTCCTCTGTTACTGGTTACAGAGT 4989  
 Db 5024 GGACAAACAGCATTAGTGTCAAGTGGCTGCTTCAAGTTCCTCTGTTACTGGTTACAGAGT 5083  
 Qy 4990 AACCAACCACTCCCAAAATATGACACAGGACCAACAAATACTAAACTGACAGTCCAGATCA 5049  
 Db 5084 AACCAACCACTCCCAAAATATGACACAGGACCAACAAATACTAAACTGACAGTCCAGATCA 5143  
 Qy 5050 AACAGAAATGACTATTGAAAGGCTTCAGCCCAACAGTGGAGTATGTTAGTGTCTATGC 5109  
 Db 5144 AACAGAAATGACTATTGAAAGGCTTCAGCCCAACAGTGGAGTATGTTAGTGTCTATGC 5203  
 Qy 5110 TCAGAAATCCAAACGAGAGAGTCAACCTCTGCTTCACTGAGTAAACCAATTTGATCG 5169  
 Db 5204 TCAGAAATCCAAACGAGAGAGTCAACCTCTGCTTCACTGAGTAAACCAATTTGATCG 5263  
 Qy 5170 CCTTAAAGGACGTGGATTCACCTGATGTGGATGTCCATTCATCAAAATTTGTTGGGAAAG 5229  
 Db 5264 CCTTAAAGGACGTGGATTCACCTGATGTGGATGTCCATTCATCAAAATTTGTTGGGAAAG 5323  
 Qy 5230 CCCACAGGGGCAAGTTTCCAGGTACAGGTTGACCTTACTCGAGCCCTGAGGATGGAAATCA 5289  
 Db 5324 CCCACAGGGGCAAGTTTCCAGGTACAGGTTGACCTTACTCGAGCCCTGAGGATGGAAATCA 5383  
 Qy 5290 TGAGCTTATTTCCCTGCACTCTGATGTGGATGTGGATGTCCATTCATCAAAATTTGTTGGGAAAG 5349  
 Db 5384 TGAGCTTATTTCCCTGCACTCTGATGTGGATGTGGATGTCCATTCATTCAGAGCTGCAAGGCTCAGACC 5443  
 Qy 5350 GGGTTCCTGAGTACACAGTCAAGTGTGGTTCCTTGCACGATGATGAGAGGACGACCCCT 5409  
 Db 5444 GGGTTCCTGAGTACACAGTCAAGTGTGGTTCCTTGCACGATGATGAGAGGACGACCCCT 5503  
 Qy 5410 GATTGGAACCCAGTCCAAGCTTATTCCTGACCACTGACCTGAAAGTTCACTCAGGTCAAC 5469  
 Db 5504 GATTGGAACCCAGTCCAAGCTTATTCCTGACCACTGACCTGAAAGTTCACTCAGGTCAAC 5563  
 Qy 5470 ACCCAACAGCTGAGCGCCAGTGGACACCACTTTCAGCTCACTGAGTATCGAGT 5529  
 Db 5564 ACCCAACAGCTGAGCGCCAGTGGACACCACTTTCAGCTCACTGAGTATCGAGT 5623  
 Qy 5530 GGGGTGACCCCAAGGAGAACCGGACCAATGAAAGAAATCAACCTTGTCTCTGACAG 5589  
 Db 5624 GGGGTGACCCCAAGGAGAACCGGACCAATGAAAGAAATCAACCTTGTCTCTGACAG 5683  
 Qy 5590 CTCATCCGCTGTTGATCAGGACTTATGGTGGCCACCAATATGAAAGTGAAGTGTCTATGC 5649  
 Db 5684 CTCATCCGCTGTTGATCAGGACTTATGGTGGCCACCAATATGAAAGTGAAGTGTCTATGC 5743  
 Qy 5650 TCTTAAAGGACACTTTGAACAGCAGACCTCAGGAGTGTGTCACCACTCTGGAGAAATGT 5709  
 Db 5744 TCTTAAAGGACACTTTGAACAGCAGACCTCAGGAGTGTGTCACCACTCTGGAGAAATGT 5803  
 Qy 5710 CAGCCCAACAGAGGCTGTCAGCAGATCTACTGAGACCACTCAGCATTAGCTG 5769  
 Db 5804 CAGCCCAACAGAGGCTGTCAGCAGATCTACTGAGACCACTCAGCATTAGCTG 5863  
 Qy 5770 GAGAACCAAGACTGAGACGATCACTGGCTTCAAGTGTGATGCGTTCAGCCAAATGGCCA 5829  
 Db 5864 GAGAACCAAGACTGAGACGATCACTGGCTTCAAGTGTGATGCGTTCAGCCAAATGGCCA 5923  
 Qy 5830 GACTCCAATCCAGAGAACCATCAGCCAGATCTCAGAGCTACACCATCAGTGGCTTACA 5889  
 Db 5924 GACTCCAATCCAGAGAACCATCAGCCAGATCTCAGAGCTACACCATCAGTGGCTTACA 5983  
 Qy 5890 ACCAGGACTGACTCAAGATCTACTGTCACCTTGAATGAAATGCTGCGAGCTCCCC 5949  
 Db 5984 ACCAGGACTGACTCAAGATCTACTGTCACCTTGAATGAAATGCTGCGAGCTCCCC 6043  
 Qy 5950 TGTGCTCATCGACGCTTCACTGCAATGATGCACTCCAACTGCGTTCCTGGCCAC 6009  
 Db 6044 TGTGCTCATCGACGCTTCACTGCAATGATGCACTCCAACTGCGTTCCTGGCCAC 6103

5010 CACACCAATTCCTTCTGGTATCATGGCAGCCGACAGTTCAGGATTACCGGCTACAT 6069  
 b 5104 CACACCAATTCCTTCTGGTATCATGGCAGCCGACAGTTCAGGATTACCGGCTACAT 6163  
 y 6070 CATCAAGTATGAGAACCTTGGTCTCTCCAGAGAAAGTGTCCCTCGGCCCGCCCTGG 6129  
 b 6164 CATCAAGTATGAGAACCTTGGTCTCTCCAGAGAAAGTGTCCCTCGGCCCGCCCTGG 6223  
 y 6130 TGTTCAGAGGCTACTATTACTTGGCTGGACCCGGGAACCGGAATATACAAATTTATGTAT 6189  
 b 6224 TGTTCAGAGGCTACTATTACTTGGCTGGACCCGGGAACCGGAATATACAAATTTATGTAT 6283  
 y 6190 TGCCCTGAAGAAATAATCAGAGAGGAGCCCTGTATTTGAAGGAAAAAGACAGAGCT 6249  
 b 6284 TGCCCTGAAGAAATAATCAGAGAGGAGCCCTGTATTTGAAGGAAAAAGACAGAGCT 6343  
 y 6250 TGCCCAACTGTTAAACCTTCCACACCCCAATCTTCATGGACAGAGATCTTGGATTTCC 6309  
 b 6344 TGCCCAACTGTTAAACCTTCCACACCCCAATCTTCATGGACAGAGATCTTGGATTTCC 6403  
 y 6310 TTCCACAGTTCAAAAGACCCCTTTGTCCACCACTGGTGGTATGACATGGAAATGGTAT 6369  
 b 6404 TTCCACAGTTCAAAAGACCCCTTTGTCCACCACTGGTGGTATGACATGGAAATGGTAT 6463  
 y 6370 TCAGCTTCTGGCACTCTGTGTGAGCAACCCAGAGTTGGGCAACAAATGATCTTTGAGGA 6429  
 b 6464 TCAGCTTCTGGCACTCTGTGTGAGCAACCCAGAGTTGGGCAACAAATGATCTTTGAGGA 6523  
 y 6430 ACATGTTTTAGCGGACACACCCGCCCAACAGGCCACCCCAATAGGCATAGGCCAAG 6489  
 b 6524 ACATGTTTTAGCGGACACACCCGCCCAACAGGCCACCCCAATAGGCATAGGCCAAG 6583  
 y 6490 ACATGTTTTAGCGGACACACCCGCCCAACAGGCCACCCCAATAGGCATAGGCCAAG 6507  
 b 6584 ACCATACCCGCGGAATGT-----AGGCAAGA 6516  
 y 6508 -----AGGCAAGA 6703  
 b 6644 AGACTATACCTGTACCCACACGTCGGGGCTCAATCCAAATGCTCTACAGGACAGA 6703  
 y 6517 AGCTCTCTCAGACAAACATCTCATGGGCCCATTCAGGACACTCTCAGTACATCAT 6576  
 b 6704 AGCTCTCTCAGACAAACATCTCATGGGCCCATTCAGGACACTCTCAGTACATCAT 6763  
 y 6577 TTCAATGTCATCTCTGGCACTCATGAAGAACCTTTACAGTTTCAGGGTTCTCGGAACCTC 6636  
 b 6764 TTCAATGTCATCTCTGGCACTCATGAAGAACCTTTACAGTTTCAGGGTTCTCGGAACCTC 6823  
 y 6637 TACAGTGCCATCTCAGACAGGCTCAGCAGAGTGCCACCTACACATCATATAGTGAGGC 6696  
 b 6824 TACAGTGCCATCTCAGACAGGCTCAGCAGAGTGCCACCTACACATCATATAGTGAGGC 6883  
 y 6697 ACTGAAGACCCAGACAGGATAGGTTTCGGGAAGAGTTGTACCGTGGCAACTCTGT 6756  
 b 6884 ACTGAAGACCCAGACAGGATAGGTTTCGGGAAGAGTTGTACCGTGGCAACTCTGT 6943  
 y 6757 CAACGAAGGTTGAACCAACCTACGAGTACTCGTGTGACCCCTACACAGTGTCCCA 6816  
 b 6944 CAACGAAGGTTGAACCAACCTACGAGTACTCGTGTGACCCCTACACAGTGTCCCA 7003  
 y 6817 TTATGCGTTTGGAGATGAGTGGGAACGAATGTCTGAATCAGGCTTTAAACTGTGTGCCA 6876  
 b 7004 TTATGCGTTTGGAGATGAGTGGGAACGAATGTCTGAATCAGGCTTTAAACTGTGTGCCA 7063  
 y 6877 GTGCTTAGGCTTTGGAAAGTGGTCAATTCAGATGTGATTCATCTAGATGGTGGCCATGACAA 6936  
 b 7064 GTGCTTAGGCTTTGGAAAGTGGTCAATTCAGATGTGATTCATCTAGATGGTGGCCATGACAA 7123  
 y 6937 TGCTGTCAACTACAAAGTTGGAGAGTGGGACCTCAGGAGGAAAAATGGCCAGATGAT 6996  
 b 7124 TGCTGTCAACTACAAAGTTGGAGAGTGGGACCTCAGGAGGAAAAATGGCCAGATGAT 7183  
 y 6997 GAGCTGCACATGTCTTGGGAACGGAAAAAGGAATTCAGGTGTGACCCCTCATGAGGCAAC 7056

7184 GAGCTGCACATGTCTTGGGAACGGAAAGAGAAATTCAGTGTGACCCCTCATGAGGCAAC 7243  
 y 7057 GTGTTACGATGATGGAGAGACATACCACTAGAGAACAGTGGCAGAAAGAAATATCTCGG 7116  
 b 7244 GTGTTACGATGATGGAGAGACATACCACTAGAGAACAGTGGCAGAAAGAAATATCTCGG 7303  
 y 7117 TGCATTTGCTCTGTGCATCATGTCTTGGAGGCGCAGCGGGCTGGCGTGTGACAACTGGC 7176  
 b 7304 TGCATTTGCTCTGTGCATCATGTCTTGGAGGCGCAGCGGGCTGGCGTGTGACAACTGGC 7363  
 y 7177 CAGACCTGGGGGTGAACCCAGTCCCGAAGGCACTACTTGGCCAGTCTACAAACAGTATTC 7236  
 b 7364 CAGACCTGGGGGTGAACCCAGTCCCGAAGGCACTACTTGGCCAGTCTACAAACAGTATTC 7423  
 y 7237 TCAGAGATACCATCAGAGAACAAACCTAATGTAAATTCGCCAATTTGAGTGTCTCATGCC 7296  
 b 7424 TCAGAGATACCATCAGAGAACAAACCTAATGTAAATTCGCCAATTTGAGTGTCTCATGCC 7483  
 y 7297 TTTAGATGTACAGGCTGACAGAGAGATTCCTGAGGAGTAAATCATCTTTCCAAATCCAGAG 7356  
 b 7484 TTTAGATGTACAGGCTGACAGAGAGATTCCTGAGGAGTAAATCATCTTTCCAAATCCAGAG 7543  
 y 7357 GAACAAGCATGTCTCTGCCAAGATCCCATCTAAACTGGAGTGTATAGCAGACCCAGC 7416  
 b 7544 GAACAAGCATGTCTCTGCCAAGATCCCATCTAAACTGGAGTGTATAGCAGACCCAGC 7603  
 y 7417 TTAGAGTTCTCTTCTTCTTAAAGCCCTTTGCTCTGAGGAGTAAATCATCTTTCCAAATCCAGAG 7476  
 b 7604 TTAGAGTTCTCTTCTTCTTAAAGCCCTTTGCTCTGAGGAGTAAATCATCTTTCCAAATCCAGAG 7663  
 y 7477 CAATCTCAGCTTCTCCTCAAGCATCACCTGGGAGTTTCCTGAGGGTTTCTCTCAATAATGA 7536  
 b 7664 CAATCTCAGCTTCTCCTCAAGCATCACCTGGGAGTTTCTCTGAGGGTTTCTCTCAATAATGA 7723  
 y 7537 GGGTGCACATTCGCTGTCTGAGTATTCATACCGCTCAGTATTTTAAATGAA 7596  
 b 7724 GGGTGCACATTCGCTGTCTGAGTATTCATACCGCTCAGTATTTTAAATGAA 7783  
 y 7597 GTGATTTCTAAGATTTGGTTTGGGATCAATAGGAAGCATATGAGCCAAACCAAGATGCAA 7656  
 b 7784 GTGATTTCTAAGATTTGGTTTGGGATCAATAGGAAGCATATGAGCCAAACCAAGATGCAA 7843  
 y 7657 ATGTTTTGAAATGATGACCAAAATTTTAAAGTAGGAAGTACCCCAACACTCTGCTT 7716  
 b 7844 ATGTTTTGAAATGATGACCAAAATTTTAAAGTAGGAAGTACCCCAACACTCTGCTT 7903  
 y 7717 TCATTAAGTGTCTGGCCCGCAATCTGTAGGAACAGCATGATCTTGTACTGTGATAT 7776  
 b 7904 TCATTAAGTGTCTGGCCCGCAATCTGTAGGAACAGCATGATCTTGTACTGTGATAT 7963  
 y 7777 TTTAAATATCCACAGTACT 7795  
 b 7964 TTTAAATATCCACAGTACT 7982

RESULT 6  
 US-10-098-841-5  
 ; Sequence 5, Application US/10098841  
 ; Publication No. US20020197679A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Liu, Chonghua  
 ; APPLICANT: Asundi, Vinod  
 ; APPLICANT: Xu, Chongjun  
 ; APPLICANT: Zhou, Ping  
 ; APPLICANT: Wang, Jiong-Rui  
 ; APPLICANT: Zhao, Qing A.  
 ; APPLICANT: Ren, Feiyan  
 ; APPLICANT: Chen, Rui-hong  
 ; APPLICANT: Wang, Dunrui  
 ; APPLICANT: Wang, Zhiwei





1750 GAGGTGACGTCGTGTTGGGAATGGTCTGGGGAATGGACATGCTACTGCGAGCT 1809  
b 1844 GAGGTGACGTCGTGTTGGGAATGGTCTGGGGAATGGACATGCTACTGCGAACT 1903  
y 1810 TCGAGATCAGTCATGTTGATGACATCACTTACATGTAAGTGAACGACATCCCAAGCG 1869  
b 1904 TCGAGATCAGTCATGTTGATGACATCACTTACATGTAAGTGAACGACATCCCAAGCG 1963  
y 1870 TCATGAAGAGGGGACATGCTGAACCTGTACATGCTTCGGTCAAGGTTCGGGCGAGTGGAA 1929  
b 1964 TCATGAAGAGGGGACATGCTGAACCTGTACATGCTTCGGTCAAGGTTCGGGCGAGTGGAA 2023  
y 1930 GTGTGATCCCGTCGACCAATGCCAGATTCAGAGCTGGAGTGGATTTATCAAAATGGAGA 1989  
b 2024 GTGTGATCCCGTCGACCAATGCCAGATTCAGAGCTGGAGTGGATTTATCAAAATGGAGA 2083  
y 1990 TTCAATGGGAGAGTATGTGATCGATGGTTCAGATACCAGTCTACTGCTATGGCCGTGGCAT 2049  
b 2084 TTCAATGGGAGAGTATGTGATGGTTCAGATACCAGTCTACTGCTATGGCCGTGGCAT 2143  
y 2050 TGGGAGTGGCATTTGCCAATTTACAGACCTTATCCAGCTCAAGTGGTCTGTGCGAAT 2109  
b 2144 TGGGAGTGGCATTTGCCAATTTACAGACCTTATCCAGCTCAAGTGGTCTGTGCGAAT 2203  
y 2110 ATTATATCACTGAGACTCCGAGTCAGGCCAACTCCACCCCAATCCAGTGGAAATGCAACCA 2169  
b 2204 ATTATATCACTGAGACTCCGAGTCAGGCCAACTCCACCCCAATCCAGTGGAAATGCAACCA 2263  
y 2170 GCCATCTCAATTTCCAAATGTAATTCAGTGGAGACCTTAAATTCCTGTGAGGCGGTG 2229  
b 2264 GCCATCTCAATTTCCAAATGTAATTCAGTGGAGACCTTAAATTCCTGTGAGGCGGTG 2323  
y 2230 GAAGGAGTCTACCATACGAGGCCACTTAACTCTACACCATCAAGAGGCTCAAGCCTGG 2289  
b 2324 GAAGGAGTCTACCATACGAGGCCACTTAACTCTACACCATCAAGAGGCTCAAGCCTGG 2383  
y 2290 TGTGATACGAGGCGAGTCATACAGTCTCAGCAGTACGCGACCGCCACCAAGAGTACTCG 2349  
b 2384 TGTGATACGAGGCGAGTCATACAGTCTCAGCAGTACGCGACCGCCACCAAGAGTACTCG 2443  
y 2350 CTATTGACTTACCAACACGACGACGACGACCTGTGACGACCAACCGCTGACAGGAGA 2409  
b 2444 CTATTGACTTACCAACACGACGACGACGACCTGTGACGACCAACCGCTGACAGGAGA 2503  
y 2410 GACGACTCCCTTTCTCTCTGTCGCACTTCTGAATCTGTGACGCAATACAGCCGAC 2469  
b 2504 GACGACTCCCTTTCTCTCTGTCGCACTTCTGAATCTGTGACGCAATACAGCCGAC 2563  
y 2470 TAGCTTTGTGGTCTCTGTTGGTCTCAGCTTCGACACCGCTGTCGGGATTCGGGTCGAATA 2529  
b 2564 TAGCTTTGTGGTCTCTGTTGGTCTCAGCTTCGACACCGCTGTCGGGATTCGGGTCGAATA 2623  
y 2530 TGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGATCTTCCAAGCACAGCCACTTC 2589  
b 2624 TGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGATCTTCCAAGCACAGCCACTTC 2683  
y 2590 TGTGAACATCCCTGACCTGCTTCTGGCGAAATACATGTTGAATGTCTATCAGATATC 2649  
b 2684 TGTGAACATCCCTGACCTGCTTCTGGCGAAATACATGTTGAATGTCTATCAGATATC 2743  
y 2650 TGAGGATGGGAGCAGAGTTGTATCTCTGCTACTTCCAAACACAGCCGCTGATGCC 2709  
b 2744 TGAGGATGGGAGCAGAGTTGTATCTCTGCTACTTCCAAACACAGCCGCTGATGCC 2803  
y 2710 TCTGACCCGACGTCGTGACCAAGTGTGATGACCTCAATGTTGTTTCGCTGGAGCAGCC 2769  
b 2804 TCTGACCCGACGTCGTGACCAAGTGTGATGACCTCAATGTTGTTTCGCTGGAGCAGCC 2863  
y 2770 CCAGGCTCCCATCACAGGGTACAGATAGTCTATTCGCCATCAGTAGAAGGTAGCAGCAC 2829  
b 2864 CCAGGCTCCCATCACAGGGTACAGATAGTCTATTCGCCATCAGTAGAAGGTAGCAGCAC 2923  
y 2830 AGAAGTCAACCTTCCTGAAAGTGGTTCGCTGATCCGCTCAGTCTGACCTGCAAGCTGGTGT 2889

2924 AGAAGTCAACCTTCCTGAAAGTGGTTCGCTGATCCCTGATGCTTCCAAACCTGGTGT 2983  
y 2890 TCAATATAACATCACTATCTATGCTGTGGAAGAAATCAAGAAAGTACACCTGTTGTCTAT 2949  
b 2984 TCAATATAACATCACTATCTATGCTGTGGAAGAAATCAAGAAAGTACACCTGTTGTCTAT 3043  
y 2950 TCAACAAAGAAACCACTTGGGACCCGCTCAGATACAGTGGCTTCCAGGAGCCTGCA 3009  
b 3044 TCAACAAAGAAACCACTTGGGACCCGCTCAGATACAGTGGCTTCCAGGAGCCTGCA 3103  
y 3010 GTTTGTGGAAGTACAGAGCTGAAGGTCAACCTATCATGTGACACACCCCTGAGAGTGCAGT 3069  
b 3104 GTTTGTGGAAGTACAGAGCTGAAGGTCAACCTATCATGTGACACACCCCTGAGAGTGCAGT 3163  
y 3070 GACCGGTACCGTGTGATGTGATCCCGTCAACCTTGGGAGCAGCGGAGAGGCT 3129  
b 3164 GACCGGTACCGTGTGATGTGATCCCGTCAACCTTGGGAGCAGCGGAGAGGCT 3223  
y 3130 GCCCATCAGCAGGAACACCTTTGACAGAGTCAACCGGCTGTCCCTTGGGCTCACCTATTATTA 3189  
b 3224 GCCCATCAGCAGGAACACCTTTGACAGAGTCAACCGGCTGTCCCTTGGGCTCACCTATTATTA 3283  
y 3190 CTTCAAGTCTTTGCAAGTGCAGTGGAGGAGAGCAAGCTCTGACTGCTCAACAGAC 3249  
b 3284 CTTCAAGTCTTTGCAAGTGCAGTGGAGGAGAGCAAGCTCTGACTGCTCAACAGAC 3343  
y 3250 AACCAGTCTGATGCTCCCACTAACTTCAAGTGTCAATGAACTGATTTCTACTGTCCT 3309  
b 3344 AACCAGTCTGATGCTCCCACTAACTTCAAGTGTCAATGAACTGATTTCTACTGTCCT 3403  
y 3310 GGTGAGTGAATCCACTCGGCGCCAGATTAACAGGATACCGACTGACCGTGGGCGCTTAC 3369  
b 3404 GGTGAGTGAATCCACTCGGCGCCAGATTAACAGGATACCGACTGACCGTGGGCGCTTAC 3463  
y 3370 CCGAAGGAGCAGCCAGGCGAGTACATGTTGGTCTCTGCTCCAAAGTACCACTGAG 3429  
b 3464 CCGAAGGAGCAGCCAGGCGAGTACATGTTGGTCTCTGCTCCAAAGTACCACTGAG 3523  
y 3430 GAATCTCAGCCTGCACTGAGTACACCGTATCCCTGTTGGCCATTAAGGGCAACCAAGA 3489  
b 3524 GAATCTCAGCCTGCACTGAGTACACCGTATCCCTGTTGGCCATTAAGGGCAACCAAGA 3583  
y 3490 GAGCCCAAGCCACTGGAGTCTTTACCACTGACGCTGGGAGTCTATTCCACCTTA 3549  
b 3584 GAGCCCAAGCCACTGGAGTCTTTACCACTGACGCTGGGAGTCTATTCCACCTTA 3643  
y 3550 CAACACCGAGTGAAGTGAACCCACTTGTGATCACAATGACCGCTGCTCCAAAGAAATGG 3609  
b 3644 CAACACCGAGTGAAGTGAACCCACTTGTGATCACAATGACCGCTGCTCCAAAGAAATGG 3703  
y 3610 TTTTAAGCTGGTGTACGACCAAGCCAGGAGGAGGACCAACGAGAGTGAAGTTCGAGA 3669  
b 3704 TTTTAAGCTGGTGTACGACCAAGCCAGGAGGAGGACCAACGAGAGTGAAGTTCGAGA 3763  
y 3670 CTCAGGAAGCATCTGTTGTCGGCTTGAATCCAGGAGTGAATACGCTTACACCTTCA 3729  
b 3764 CTCAGGAAGCATCTGTTGTCGGCTTGAATCCAGGAGTGAATACGCTTACACCTTCA 3823  
y 3730 AGTCTGTGAGATGAGCAGGAAGAGATGGCCAAATTTGTAAACAAAGTGTGACCACTT 3789  
b 3824 AGTCTGTGAGATGAGCAGGAAGAGATGGCCAAATTTGTAAACAAAGTGTGACCACTT 3883  
y 3790 GTCTCCCAACCAAACTTGCATCTGGAGGCAAAACCTTGACACTGGAGTGTCTCAGTCTC 3849  
b 3884 GTCTCCCAACCAAACTTGCATCTGGAGGCAAAACCTTGACACTGGAGTGTCTCAGTCTC 3943  
y 3850 CTGGAGAGGAGCAGCAACCCCGACATTTACTGTTATAGAAATACCAACACCCCTACAAA 3909  
b 3944 CTGGAGAGGAGCAGCAACCCCGACATTTACTGTTATAGAAATACCAACACCCCTACAAA 4003  
y 3910 CGGCCAGCAGGGAATTTCTTTGGAGAAAGTGGTCCATGCTGATCAGAGCTCTCTGCACTTT 3969





6130 TGTCAAGAGGCTACTATTACTGGCTGGAAACGGGAACCGGAATATACAATTTATGTCAAT 6189  
6224 TGTCAAGAGGCTACTATTACTGGCTGGAAACGGGAACCGGAATATACAATTTATGTCAAT 6283  
6190 TGGCCCTGAAGATATATCAAGAGGCGAGCCCTGTATTTGAAGGAAGAAAGACAGACGAGCT 6249  
6284 TGGCCCTGAAGATATATCAAGAGGCGAGCCCTGTATTTGAAGGAAGAAAG----- 6332  
6250 TCCCCAACTGGTAACCCCTTCCACACCCCAATCTTCATGGACAGAGATCTTTGGATGTTTCC 6309  
6333 ----- 6332  
6310 TTCCACAGTTCAAAAGACCCCTTTCTGTACCCACCCCTGGGTATGACACTGGAAATGGTAT 6369  
6333 -----ACAGTTCAAAAGACCCCTTTCTGTACCCACCCCTGGGTATGACACTGGAAATGGTAT 6388  
6370 TCAGCTTCTCTGGCACTTCTGGTTCAGCAACCCCAAGTGTGGGCAACCAATGATCTTTGAGGA 6429  
6389 TCAGCTTCTCTGGCACTTCTGGTTCAGCAACCCCAAGTGTGGGCAACCAATGATCTTTGAGGA 6448  
6430 ACATGGTTTTAGCGGACACACACCGCCCAAGCGGCCACCCCATATAGGCATAGGCCAAG 6489  
6449 ACATGGTTTTAGCGGACACACACCGCCCAAGCGGCCACCCCATATAGGCATAGGCCAAG 6508  
6490 ACCATACCCGCGCAATGTAGGACCAAGAGCTCTCTCAGACAAACCATCTCATAGGCGCCC 6549  
6509 ACCATACCCGCGCAATGTAGGACCAAGAGCTCTCTCAGACAAACCATCTCATAGGCGCCC 6568  
6550 ATTCCAGGACACTTCTGAGTACATCAATTTCAATGTCATCTCTGTGGCACTGATGAAGACC 6609  
6569 ATTCCAGGACACTTCTGAGTACATCAATTTCAATGTCATCTCTGTGGCACTGATGAAGACC 6628  
6610 CTTACAGTTTCAGGGTTCTTGGAACTTCTACCAAGTGCCTCTGACAGGCTCTCACCAGAGG 6669  
6629 CTTACAGTTTCAGGGTTCTTGGAACTTCTACCAAGTGCCTCTGACAGGCTCTCACCAGAGG 6688  
6670 TGGCACTCAACATCATAGTGGAGGACATGAAGACCAAGGAGGACATAGGTTGCGGA 6729  
6689 TGGCACTCAACATCATAGTGGAGGACATGAAGACCAAGGAGGACATAGGTTGCGGA 6748  
6730 AGAGGTTGTACCGTGGGCACTCTCTCAACAGGCTTGAACCAACCTACCGATGACTC 6789  
6749 AGAGGTTGTACCGTGGGCACTCTCTCAACAGGCTTGAACCAACCTACCGATGACTC 6808  
6790 GTCTTTGACCCCTACACAGTGTCCCATATGACCGTTGAGATGAGTGGGAACGAATGTC 6849  
6809 GTCTTTGACCCCTACACAGTGTCCCATATGACCGTTGAGATGAGTGGGAACGAATGTC 6868  
6850 TGAATCAGGCTTTAAACTGTTGCGCAGTGTCTGAGGCTTTGGAAGTGGTCATTTCAGATG 6909  
6869 TGAATCAGGCTTTAAACTGTTGCGCAGTGTCTGAGGCTTTGGAAGTGGTCATTTCAGATG 6928  
6910 TGATTCATCTAGATGGTGCATCACAATGGTGTGAATACTAAGATTTGGAGAGAGTGGGA 6969  
6929 TGATTCATCTAGATGGTGCATCACAATGGTGTGAATACTAAGATTTGGAGAGAGTGGGA 6988  
6970 CCGTCAGGAGGAATAGCGCAGATGATGAGTGCACATGCTTTGGAAACGGAAGAGGA 7029  
6989 CCGTCAGGAGGAATAGCGCAGATGATGAGTGCACATGCTTTGGAAACGGAAGAGGA 7048  
7030 ATTCAAGTGTGACCCCTCATGAGGCAACGTTGTTACGATGATGGGAAGACATACCAAGTAGG 7089  
7049 ATTCAAGTGTGACCCCTCATGAGGCAACGTTGTTACGATGATGGGAAGACATACCAAGTAGG 7108  
7090 AGAACAGTGCAGAGGAATATCTCGTGCAATTTGCTCTGACATGCTTTGGAGGCA 7149  
7109 AGAACAGTGCAGAGGAATATCTCGTGCAATTTGCTCTGACATGCTTTGGAGGCA 7168  
7150 GCGGGGCTGGCGTGTGACAACTGCCGACACCTGGGGGTGAACCCAGTCCCGAAGGAC 7209  
7169 GCGGGGCTGGCGTGTGACAACTGCCGACACCTGGGGGTGAACCCAGTCCCGAAGGAC 7228  
7210 TACTGGCCAGTCTCTACAAACCAAGTATTTCTCAGAGATACCATCAGAGAACAAACACTAATGT 7269

Db 7229 TACTGGCCAGTCTCTACAACAGATTTCTCAGAGATACCATCAGAGAACAAACACTAATGT 7288  
Qy 7270 TAATTTGCCCAATTAAGTGTCTTTCATGCTTTAGATGTACAGGCTGACAGAGAAGATTCGG 7329  
Db 7289 TAATTTGCCCAATTAAGTGTCTTTCATGCTTTAGATGTACAGGCTGACAGAGAAGATTCGG 7348  
Qy 7330 AGAGTAATATCATCTTTCCCAATCCAGAGGACAGAGCATGTCTCTGCGCAAGATCCATCTA 7389  
Db 7349 AGAGTAATATCATCTTTCCCAATCCAGAGGACAGAGCATGTCTCTGCGCAAGATCCATCTA 7408  
Qy 7390 AACTGGAGTCAATGTTAGCAGACCCAGCTTAGAGTCTTCTTCTTTTAAAGCCCTTTGC 7449  
Db 7409 AACTGGAGTCAATGTTAGCAGACCCAGCTTAGAGTCTTCTTCTTTTAAAGCCCTTTGC 7468  
Qy 7450 TCTGGAGGAATGTTCTCCAGCTTCAGCTCAACTCAGAGCTTCTCCAGAGCATCACCCTGGGA 7509  
Db 7469 TCTGGAGGAATGTTCTCCAGCTTCAGCTCAACTCAGAGCTTCTCCAGAGCATCACCCTGGGA 7528  
Qy 7510 GTTTCCTGAGGGTTTTCTCATAAATGAGGCTGCACATTTGCCCTGTTCTGCTTCGAAGTAT 7569  
Db 7529 GTTTCCTGAGGGTTTTCTCATAAATGAGGCTGCACATTTGCCCTGTTCTGCTTCGAAGTAT 7588  
Qy 7570 TCAATACCGTCTAGTATTTTAAATGAAGTGAATTTGAAGATTTGGTTGGATCAATAGGA 7629  
Db 7589 TCAATACCGTCTAGTATTTTAAATGAAGTGAATTTGAAGATTTGGTTGGATCAATAGGA 7648  
Qy 7630 AAGCATATGAGCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAATTTTAAAT 7689  
Db 7649 AAGCATATGAGCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAATTTTAAAT 7708  
Qy 7690 AGGAAGTCAACCAACACATTTCTGCTTTCACTTAAGTGTCTGCGCCGCAATACTGTAGGA 7749  
Db 7709 AGGAAGTCAACCAACACATTTCTGCTTTCACTTAAGTGTCTGCGCCGCAATACTGTAGGA 7768  
Qy 7750 ACAGAGTATCTTTGTTACTGTATATTTTAAATATCCACAGTACT 7795  
Db 7769 ACAAGATATCTTTGTTACTGTATATTTTAAATATCCACAGTACT 7814

RESULT 7

US-09-964-824A-574  
; Sequence 574, Application US/09964824A  
; Patent No. US20020102531A1  
; GENERAL INFORMATION:  
; APPLICANT: Horrigan, Stephen  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
; FILE OF INVENTION: Sets  
; FILE REFERENCE: 689290-73  
; CURRENT APPLICATION NUMBER: US/09/964,824A  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US/60/236,033  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,032  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,028  
; PRIOR FILING DATE: 2000-09-28  
; NUMBER OF SEQ ID NOS: 583  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 574  
; LENGTH: 7680  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-964-824A-574

Query Match 95.2%; Score 7423; DB 9; Length 7680;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
Qy 349 GAAGAGCAAGAGGAGGAGGCTCAGCAAAATGGTTTCAGCCCAAGTCCCGGTGGCTGTCACTCA 408  
Db 1 GAAGAGCAAGAGGAGGAGGCTCAGCAAAATGGTTTCAGCCCAAGTCCCGGTGGCTGTCACTCA 60

QY 409 AAGCAAGCCCGTTGTTATGCAATGGAAACACATATCAGATAAATCAACAGTGGGAGCG 468  
DB 61 AAGCAAGCCCGTTGTTATGCAATGGAAACACATATCAGATAAATCAACAGTGGGAGCG 120  
QY 469 CACCTACCTAGGCAATCGTTGGTTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTG 528  
DB 121 GACCTACCTAGGTAATGTGTTGGTTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTG 180  
QY 529 CGAAGTAACCTGAGCTGAAGAGACTTGTCTTTGACAGATGACTGGGAACACTTACCG 588  
DB 181 CGAAGTAACCTGAGCTGAAGAGACTTGTCTTTGACAGATGACTGGGAACACTTACCG 240  
QY 589 AGTGGGTGACACTTATGAGCGTCTTAAGAATCCATGATCTGGGACTGTATCTGCACTCGG 648  
DB 241 AGTGGGTGACACTTATGAGCGTCTTAAGAATCCATGATCTGGGACTGTATCTGCACTCGG 300  
QY 649 GGCTGGGGGAGGAGATAAAGCTGTACCATCGAAGCCGCTGCCATGAAGGGGTGAGTC 708  
DB 301 GGCTGGGGGAGGAGATAAAGCTGTACCATCGAAGCCGCTGCCATGAAGGGGTGAGTC 360  
QY 709 CTACAAGATTGGTGACCTGGAGGAGACCATGAGACTGGTGGTTACATGTTAGAGTG 768  
DB 361 CTACAAGATTGGTGACCTGGAGGAGACCATGAGACTGGTGGTTACATGTTAGAGTG 420  
QY 769 TGTGTGCTTGGTAATGAAGAGAAATGGAACCTGCAAGCCCATAGCTGAGAAAGTGT 828  
DB 421 TGTGTGCTTGGTAATGAAGAGAAATGGAACCTGCAAGCCCATAGCTGAGAAAGTGT 480  
QY 829 TGATCATGCTGCTGGGACTTCTTATGTGGTGGAGAAACGTTGGAGAACCCCTACCAAGG 888  
DB 481 TGATCATGCTGCTGGGACTTCTTATGTGGTGGAGAAACGTTGGAGAACCCCTACCAAGG 540  
QY 889 CTGGATGATGTTAGATTGTAATCTGGGAGAAAGCGAGCGGACGCACTACCTGCACTTC 948  
DB 541 CTGGATGATGTTAGATTGTAATCTGGGAGAAAGCGAGCGGACGCACTACCTGCACTTC 600  
QY 949 TAGAAATAGATGCAACGATCAGGACACAAGGACATCTTATAGAAATGGAGACACTGGAG 1008  
DB 601 TAGAAATAGATGCAACGATCAGGACACAAGGACATCTTATAGAAATGGAGACACTGGAG 560  
QY 1009 CAAGAAGGATAATCAGGAAACCTGCTCAGTGGCATCTGCAAGGCAACCGCCGAGGAGA 1068  
DB 661 CAAGAAGGATAATCAGGAAACCTGCTCAGTGGCATCTGCAAGGCAACCGCCGAGGAGA 720  
QY 1069 GTGGAGTGTGAGGACACCTCTGTGAGACCAACATCGAGCGGATCTGGCCCTTCAC 1128  
DB 721 GTGGAGTGTGAGGACACCTCTGTGAGACCAACATCGAGCGGATCTGGCCCTTCAC 780  
QY 1129 CGATGTTCTGTGAGGCTGTTTACCAACCGCAGCCTCACCCCGAGCCTCTCCCTATGGCCA 1188  
DB 781 CGATGTTCTGTGAGGCTGTTTACCAACCGCAGCCTCACCCCGAGCCTCTCCCTATGGCCA 840  
QY 1189 CTGTGTACAGACAGTGTGTGTTACTCTGTGGGATGAGTGGCTGAAGACACAGG 1248  
DB 841 CTGTGTACAGACAGTGTGTGTTACTCTGTGGGATGAGTGGCTGAAGACACAGG 900  
QY 1249 AAATAAGCAATATGCTTGTGACGTGCTGGCAACCGAGTCAGCTGCCAAGAGACAGCTGT 1308  
DB 901 AAATAAGCAATATGCTTGTGACGTGCTGGCAACCGAGTCAGCTGCCAAGAGACAGCTGT 960  
QY 1309 AACCCAGACTTACGTTGGCAATCAAAATGGAGAGCCATGTGTCTTACCAATTCACCTACAA 1368  
DB 961 AACCCAGACTTACGTTGGCAATCAAAATGGAGAGCCATGTGTCTTACCAATTCACCTACAA 1020  
QY 1369 TGGCAGGAGGTTTACTCTCTGCACCAAGAGGGGACAGACGAGACATCTTTGGTGAG 1428  
DB 1021 TGGCAGGAGGTTTACTCTCTGCACCAAGAGGGGACAGACGAGACATCTTTGGTGAG 1080  
QY 1429 CACAACTTGAATATGACAGGACCAAGAAATATCTTTCTGCAAGACCAACACTGTTTT 1488  
DB 1081 CACAACTTGAATATGACAGGACCAAGAAATATCTTTCTGCAAGACCAACACTGTTTT 1140  
QY 1489 GGTTCAGACTCGAGGAGGAAATCCAAATGGTGCTTGTGGCCACTTCCCTTATACAA 1548

DB 1141 GGTTCAGACTCAAGGAGGAAATCCAAATGGTGCTTGTGGCCACTTCCCTTCTTATACAA 1200  
QY 1549 CAAACAATTAACATGATTGCACTTTCTGAGGCGAAGAGACAAACATGAGTGGTGTGG 1608  
DB 1201 CAAACAATTAACATGATTGCACTTTCTGAGGCGAAGAGACAAACATGAGTGGTGTGG 1260  
QY 1609 GACCAACAGAAACTATGATGCCGACAGAAATTTGGGTTCTGGCCCATGGGTGCCACGA 1668  
DB 1261 GACCAACAGAAACTATGATGCCGACAGAAATTTGGGTTCTGGCCCATGGGTGCCACGA 1320  
QY 1669 GGAATCTGCAAAACAATGAAGGGGTCAATGATCCGCAATGGAGATCAGTGGGATAAGCA 1728  
DB 1321 GGAATCTGCAAAACAATGAAGGGGTCAATGATCCGCAATGGAGATCAGTGGGATAAGCA 1380  
QY 1729 GATGACATGGGTCAATGATGAGGTGACGTGTGTTGGGAATGCTGTGGGGAATGAC 1788  
DB 1381 GATGACATGGGTCAATGATGAGGTGACGTGTGTTGGGAATGCTGTGGGGAATGAC 1440  
QY 1789 ATGCAATGGCTACTCGCAGCTTCGAGATGATGTCATGTTGATGACATCACTTACAATGT 1848  
DB 1441 ATGCAATGGCTACTCGCAGCTTCGAGATGATGTCATGTTGATGACATCACTTACAATGT 1500  
QY 1849 GAACGACATTCACCAAGCGTCATGAAGAGGGGACATGCTGTAACCTGTACATGCTTGG 1908  
DB 1501 GAACGACATTCACCAAGCGTCATGAAGAGGGGACATGCTGTAACCTGTACATGCTTGG 1560  
QY 1909 TCAGGGTCCGGGCGAGGTGGAAGTGTATCCCTCGAACCAATGCCAGGATTCAGAGACTGG 1968  
DB 1561 TCAGGGTCCGGGCGAGGTGGAAGTGTATCCCTCGAACCAATGCCAGGATTCAGAGACTGG 1620  
QY 1969 GAGCTTTATCAAAATGGAGATTCATGGAGAAAGTATGTCATGCTGTCAGATACAGTG 2028  
DB 1621 GAGCTTTATCAAAATGGAGATTCATGGAGAAAGTATGTCATGCTGTCAGATACAGTG 1680  
QY 2029 CTACTCTATGGCCGTGGCAATTCGGGAGTGGCAATTCGCCAACCTTTACAGACCTATCCAAG 2088  
DB 1681 CTACTCTATGGCCGTGGCAATTCGGGAGTGGCAATTCGCCAACCTTTACAGACCTATCCAAG 1740  
QY 2089 CTCAAGTGTCTGTGCGAAGTATTTATCACTGAGACTCCGAGTCCAGCCCACTCCACCC 2148  
DB 1741 CTCAAGTGTCTGTGCGAAGTATTTATCACTGAGACTCCGAGTCCAGCCCACTCCACCC 1800  
QY 2149 CATCCAGTGGAAATGACACACAGCCATCTCACAATTTCCAAAGTACATTTCTCAGGTGGAGACC 2208  
DB 1801 CATCCAGTGGAAATGACACACAGCCATCTCACAATTTCCAAAGTACATTTCTCAGGTGGAGACC 1860  
QY 2209 TAAAAATTTCTAGGCCGTTGGAAGGAAGCTACCATACAGGCCACTTAAACTCTTACAC 2268  
DB 1861 TAAAAATTTCTAGGCCGTTGGAAGGAAGCTACCATACAGGCCACTTAAACTCTTACAC 1920  
QY 2269 CATCAAGGCCCTGAAGCCTGGTGTGTATACGAGGGCCAGCTCATCAGCATCCAGCAGTA 2328  
DB 1921 CATCAAGGCCCTGAAGCCTGGTGTGTATACGAGGGCCAGCTCATCAGCATCCAGCAGTA 1980  
QY 2329 CGGCCCAAGAAAGTGAATCGCTTTCAGTTTCAACCAACAGCAACAGCAACCTGTGAC 2388  
DB 1981 CGGCCCAAGAAAGTGAATCGCTTTCAGTTTTCAGTTTCAACCAACAGCAACAGCAACCTGTGAC 2040  
QY 2389 CAGCAACCGGTGACAGGAGAGACGACTCCCTTTCTCTCTCTTGTGGCCACTTCTGAATC 2448  
DB 2041 CAGCAACCGGTGACAGGAGAGACGACTCCCTTTCTCTCTCTTGTGGCCACTTCTGAATC 2100  
QY 2449 TGTGACCGAAATCACAAGCCAGTAGCTTTGTGTCTCTCTGGGTCTCAGCTTCCGACACCGT 2508  
DB 2101 TGTGACCGAAATCACAAGCCAGTAGCTTTGTGTCTCTCTGGGTCTCAGCTTCCGACACCGT 2160  
QY 2509 GTCCGGAATTCGGGTGGAATATGAGCTGTGAGTGGGAGGAGATGAGCCACAGTACTGA 2568  
DB 2161 GTCCGGAATTCGGGTGGAATATGAGCTGTGAGTGGGAGGAGATGAGCCACAGTACTGA 2220  
QY 2569 TCTTCCAAAGCACAAGCCACTTCTGTGAACATCCCTGACCTGCTTCTGTGGCCGAAATACAT 2628

b 2221 TCTTCAAGCAGCAGCCTTCTGTGAACATCCCTGACCTGCTTCTCGGCCGAAATACAT 2280  
 y 2629 TGTAAATGTCTATACAGATATCTGAGATGGGAGCAGAGTGTGATTCCTGTCTATCTTCA 2688  
 b 2281 TGTAAATGTCTATACAGATATCTGAGATGGGAGCAGAGTGTGATTCCTGTCTATCTTCA 2340  
 y 2689 AACAAACAGCGCTGATGCGCCCTCTGACCGGACTGTGGACCAAGTTGATGACACCTCAAT 2748  
 b 2341 AACAAACAGCGCTGATGCGCCCTCTGACCGGACTGTGGACCAAGTTGATGACACCTCAAT 2400  
 y 2749 TGTGTTTCGCTGGAGCAGACCCGAGGCTCCCATACAGGGTACAGATAGTCTTATCGCC 2808  
 b 2401 TGTGTTTCGCTGGAGCAGACCCGAGGCTCCCATACAGGGTACAGATAGTCTTATCGCC 2460  
 y 2809 ATCAGTAGAAGGTAGCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCCCT 2868  
 b 2461 ATCAGTAGAAGGTAGCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCCCT 2520  
 y 2869 CAGTGACTTGCACCTGGTGTTCAGTATACATCACTATCTATGCTGTGGAAGAAATCA 2928  
 b 2521 CAGTGACTTGCACCTGGTGTTCAGTATACATCACTATCTATGCTGTGGAAGAAATCA 2580  
 y 2929 AGAAGTACACCTGTTGATTCATCAAGAAACCACTGGCACCCGCTCAGATACAGT 2988  
 b 2581 AGAAGTACACCTGTTGATTCATCAAGAAACCACTGGCACCCGCTCAGATACAGT 2640  
 y 2989 GCCCTCTCCAGGAGCCTGCGAGTTTGTGAAGTGACAGACGTTGAAGGTCAACCATCATGTG 3048  
 b 2641 GCCCTCTCCAGGAGCCTGCGAGTTTGTGAAGTGACAGACGTTGAAGGTCAACCATCATGTG 2700  
 y 3049 GACACCGCTGAGAGTGCAGTGACCGCTACCGTGTGGATGTGATCCCGTCAACCTGCC 3108  
 b 2701 GACACCGCTGAGAGTGCAGTGACCGCTACCGTGTGGATGTGATCCCGTCAACCTGCC 2760  
 y 3109 TGGCGAGCAGGCGCAGAGGTGCGCATCAGCAGGAAACACTTTGCAAGAGTCAACCGGGT 3168  
 b 2761 TGGCGAGCAGGCGCAGAGGTGCGCATCAGCAGGAAACACTTTGCAAGAGTCAACCGGGT 2820  
 y 3169 GTCCCTGGGTCACTATTAATCTTCAAGTCTTTCAGTGGCCATGGAGGAGGACAA 3228  
 b 2821 GTCCCTGGGTCACTATTAATCTTCAAGTCTTTCAGTGGCCATGGAGGAGGAGCAA 2880  
 y 3229 GCCTCTGACTGTCTCAACAGAACCAAACTGGATGCTCCCACTTAACCTCCAGTTTGTCAA 3288  
 b 2881 GCCTCTGACTGTCTCAACAGAACCAAACTGGATGCTCCCACTTAACCTCCAGTTTGTCAA 2940  
 y 3289 TGAACATGATTTACTGTCTGTGTGATGAGTCTCCACCTCGGGCCAGATTAACAGGATA 3348  
 b 2941 TGAACATGATTTACTGTCTGTGTGATGAGTCTCCACCTCGGGCCAGATTAACAGGATA 3000  
 y 3349 CCGACTGACCGTGGGCTTACCCGAGAGAGCAGCAGCCAGCAGTACAATGTGGTCCCTC 3408  
 b 3001 CCGACTGACCGTGGGCTTACCCGAGAGAGCAGCAGCCAGCAGTACAATGTGGTCCCTC 3060  
 y 3409 TGTCTCCAGTACCCACTGAGGAATCTGAGCCTGATCTGAGTACACCGTATCCCTGT 3468  
 b 3061 TGTCTCCAGTACCCACTGAGGAATCTGAGCCTGATCTGAGTACACCGTATCCCTGT 3120  
 y 3469 GGCCTAAAGGGCAACCAAGAGAGCCCAAGCCACTGGAGTCTTTACCACTGCGAGCC 3528  
 b 3121 GGCCTAAAGGGCAACCAAGAGAGCCCAAGCCACTGGAGTCTTTACCACTGCGAGCC 3180  
 y 3529 TGGGAGCTCTATTCCACCTTAAACACCGAGGTGACTGAGACCAACATTTGATACATG 3588  
 b 3181 TGGGAGCTCTATTCCACCTTAAACACCGAGGTGACTGAGACCAACATTTGATACATG 3240  
 y 3589 GAGCCTGTCTCAAGATTTGTTTAAAGTGGGTGTACGACCAAGCCAGGAGGAGGAGCC 3648  
 b 3241 GAGCCTGTCTCAAGATTTGTTTAAAGTGGGTGTACGACCAAGCCAGGAGGAGGAGCC 3300  
 y 3649 ACCACGAGAGTGTGACTTTCAGACTCAGGAAGCATCGTTGTGTCGGGTGACTCCAGGAGT 3708  
 b 3301 ACCACGAGAGTGTGACTTTCAGACTCAGGAAGCATCGTTGTGTCGGGTGACTCCAGGAGT 3360

Qy 3709 AGAATAGCTTACACCATCCAACTCCTGAGAGATGACAGGAAAGATGCGCAATTTGT 3768  
 Db 3361 AGAATAGCTTACACCATCCAACTCCTGAGATGACAGGAAAGATGCGCAATTTGT 3420  
 Qy 3769 AAACAAAGTGGTGACACCATTTCTCCAAACAACTTGCATCTGGAGGCAAAACCTTGA 3828  
 Db 3421 AAACAAAGTGGTGACACCATTTCTCCAAACAACTTGCATCTGGAGGCAAAACCTTGA 3480  
 Qy 3829 CACTGGAGTCTCACAGTCTCCTGGGAGGAGGACACACCCAGACATTACTTGGTTATAG 3888  
 Db 3481 CACTGGAGTCTCACAGTCTCCTGGGAGGAGGACACACCCAGACATTACTTGGTTATAG 3540  
 Qy 3889 AATTACCAACCCCTACAAACGCGCAGCAGGAGAAATTTCTTGGAAAGAGTGTGTCATGC 3948  
 Db 3541 AATTACCAACCCCTACAAACGCGCAGCAGGAGAAATTTCTTGGAAAGAGTGTGTCATGC 3600  
 Qy 3949 TGATCAGAGTCTCTGACATTTTGTATTAACCTGAGTCCCGGCTGGAGTACAATGTCAAGTGT 4008  
 Db 3601 TGATCAGAGTCTCTGACATTTTGTATTAACCTGAGTCCCGGCTGGAGTACAATGTCAAGTGT 3660  
 Qy 4009 TTACACTGTCAAGGATGACAAAGGAAAGTGTCCCTATCTCTGTATACCATCATCCAGCTGT 4068  
 Db 3661 TTACACTGTCAAGGATGACAAAGGAAAGTGTCCCTATCTCTGTATACCATCATCCAGCTGT 3720  
 Qy 4069 TCCTTCCTCCACTGACCTGCGATTCAACCAATTTGTCAGACCAATGCGTGTCACTTG 4128  
 Db 3721 TCCTTCCTCCACTGACCTGCGATTCAACCAATTTGTCAGACCAATGCGTGTCACTTG 3780  
 Qy 4129 GGCTCCACCCCTCCACTGATTGATTAAACCACTTCTGTCGCTTACTCACCTGTGAAAAA 4188  
 Db 3781 GGCTCCACCCCTCCACTGATTGATTAAACCACTTCTGTCGCTTACTCACCTGTGAAAAA 3840  
 Qy 4189 TGAGGAAGATGTTGACAGATGTTCAATTTCTCTTCAAGCAATGCGAGTGTCTTAACAA 4248  
 Db 3841 TGAGGAAGATGTTGACAGATGTTCAATTTCTCTTCAAGCAATGCGAGTGTCTTAACAA 3900  
 Qy 4249 TCTCCTCCCTGGTACAGATATGTAGTGTGTCTCCAGTGTCTACGAAACAAATGAGAG 4308  
 Db 3901 TCTCCTCCCTGGTACAGATATGTAGTGTGTCTCCAGTGTCTACGAAACAAATGAGAG 3960  
 Qy 4309 CACACCTCTTAGAGGAAGACAGAAACAGGTCTTGATTTCCCACTGGGATTTGACTTTTC 4368  
 Db 3961 CACACCTCTTAGAGGAAGACAGAAACAGGTCTTGATTTCCCACTGGGATTTGACTTTTC 4020  
 Qy 4369 TGATATTACTGCCAACTCTTTTACTGTGCACTGATGTCCTCGAGCCACCATCACTGG 4428  
 Db 4021 TGATATTACTGCCAACTCTTTTACTGTGCACTGATGTCCTCGAGCCACCATCACTGG 4080  
 Qy 4429 CTACAGGATCGCCATCATCCGAGCACTTTCAGTGGGAGACCTCGAGAAAGATCGGTTGCC 4488  
 Db 4081 CTACAGGATCGCCATCATCCGAGCACTTTCAGTGGGAGACCTCGAGAAAGATCGGTTGCC 4140  
 Qy 4489 CCACCTCGGAATTTCCATCACCTTCAACCACTCAGGAGGAGTATGTTGGTTCAG 4548  
 Db 4141 CCACCTCGGAATTTCCATCACCTTCAACCACTCAGGAGGAGTATGTTGGTTCAG 4200  
 Qy 4549 CATCGTTGCTCTTAATGGCAGAGGAGAAAGTCCCTTATGATTTGGCCAAACAAATCAACAGT 4608  
 Db 4201 CATCGTTGCTCTTAATGGCAGAGGAGAAAGTCCCTTATGATTTGGCCAAACAAATCAACAGT 4260  
 Qy 4609 TTCTGATGTTCCGAGGAGACCTGGAGTGTGTCGACCCCGCCAGCCCTACTGTATCAG 4668  
 Db 4261 TTCTGATGTTCCGAGGAGACCTGGAGTGTGTCGACCCCGCCAGCCCTACTGTATCAG 4320  
 Qy 4669 CTGGGATGCTCTCTCTGTCAAGTGTGATATTACAGGATCACTTACGAGAAACAGAGG 4728  
 Db 4321 CTGGGATGCTCTCTCTGTCAAGTGTGATATTACAGGATCACTTACGAGAAACAGAGG 4380  
 Qy 4729 AAATAGCCCTGTCCAGAGTTCATGTGCTGGAGGCAAGTCTTACAGTACCATTCAGCG 4788  
 Db 4381 AAATAGCCCTGTCCAGAGTTCATGTGCTGGAGGCAAGTCTTACAGTACCATTCAGCG 4440

4789 CTTTAAACCTGGAGTGTGATTATATACCTATCACTGTGTATGCTGTCACTGGCCGTGGACAG 4848  
4841 CTTTAAACCTGGAGTGTGATTATATACCTATCACTGTGTATGCTGTCACTGGCCGTGGACAG 4500  
4849 CCCCGCAAGCAGCAAGCAATTTCCATTAAATACCGAACAGAAATTGACAAAACCATCCCA 4908  
4501 CCCCGCAAGCAGCAAGCAATTTCCATTAAATACCGAACAGAAATTGACAAAACCATCCCA 4560  
4909 GATCCAGTGAACCATGTTGAGGACACAGCATTAAGTGTCAAGTGGCTGCTTCAGTTTC 4968  
4561 GATCCAGTGAACCATGTTGAGGACACAGCATTAAGTGTCAAGTGGCTGCTTCAGTTTC 4620  
4969 CCGTGTACTGTTTACAGAGTAACACCACTCCCAAAATGCAAGGCTTGCAGCCCAAGAAAC 5028  
4621 CCGTGTACTGTTTACAGAGTAACACCACTCCCAAAATGCAAGGCTTGCAGCCCAAGAAAC 4680  
5029 TAAACCTGCAGGTCAGATCAACAGAAATGACTATTGAAGGCTTGACGCCACAGTGA 5088  
4681 TAAACCTGCAGGTCAGATCAACAGAAATGACTATTGAAGGCTTGACGCCACAGTGA 4740  
5089 GTATGTGGTGTAGTGTCTATGCTCAGAAATCCAAAGGAGAGTCAAGCTCTGGTTCAGAC 5148  
4741 GTATGTGGTGTAGTGTCTATGCTCAGAAATCCAAAGGAGAGTCAAGCTCTGGTTCAGAC 4800  
5149 TGCAGTAACCAACATTTGATGCTCCCTAAAGACTGGCATTTCACTGATGTGATGTCGATTC 5208  
4801 TGCAGTAACCAACATTTGATGCTCCCTAAAGACTGGCATTTCACTGATGTGATGTCGATTC 4860  
5209 CATCAAAATGCTTGGGAAGCCCAAGGAGGCAAGTTTCCAGGTACAGGCTGACCTACTC 5268  
4861 CATCAAAATGCTTGGGAAGCCCAAGGAGGCAAGTTTCCAGGTACAGGCTGACCTACTC 4920  
5269 GAGCCCTGAGGATGGAATCCATGAGCTATTTCCTGACCTGATGTTGAAAGACACTGC 5328  
4921 GAGCCCTGAGGATGGAATCCATGAGCTATTTCCTGACCTGATGTTGAAAGACACTGC 4980  
5329 AGAGCTCAGAGGCTCAGACCGGCTCTGAGTACACAGTCAAGTGTGGTTCCTGACAGCA 5388  
4981 AGAGCTCAGAGGCTCAGACCGGCTCTGAGTACACAGTCAAGTGTGGTTCCTGACAGCA 5040  
5389 TGAATATGAGAGCCAGCCCTGATTTGAAACCCAGTCCACAGCTATTTCCTGACACCACTGA 5448  
5041 TGAATATGAGAGCCAGCCCTGATTTGAAACCCAGTCCACAGCTATTTCCTGACACCACTGA 5100  
5449 CCTGAGTTTCACTCAGTTCACACCAAGCTCAGCCGCTGAGTGAACACCACTGAT 5508  
5101 CCTGAGTTTCACTCAGTTCACACCAAGCTCAGCCGCTGAGTGAACACCACTGAT 5160  
5509 TCAGTCACTGGAATCGAGTGCAGGCTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5568  
5161 TCAGTCACTGGAATCGAGTGCAGGCTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5220  
5569 AATCAACCTTCTCAGTTCACACCAAGCTCAGCCGCTGAGTGAACACCACTGAT 5628  
5221 AATCAACCTTCTCAGTTCACACCAAGCTCAGCCGCTGAGTGAACACCACTGAT 5280  
5629 ATATGAGTGTAGTGTCTATGCTCTTAAGGACACTTTGACAGCAGACCACTCAGGAGT 5688  
5281 ATATGAGTGTAGTGTCTATGCTCTTAAGGACACTTTGACAGCAGACCACTCAGGAGT 5340  
5689 TGTCAACACTCTGAGAAATGTCAGCCCAACCAAGAGGCTGTGTGACAGATGCTACTGA 5748  
5341 TGTCAACACTCTGAGAAATGTCAGCCCAACCAAGAGGCTGTGTGACAGATGCTACTGA 5400  
5749 GACCACCATCACTTGTGAGAGCAACCAAGCTGAGACGATCACTGGCTTCCAAAGTTGA 5808  
5401 GACCACCATCACTTGTGAGAGCAACCAAGCTGAGACGATCACTGGCTTCCAAAGTTGA 5460  
5809 TGCCGTTCCAGCAATGTCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTCAGAG 5868  
5461 TGCCGTTCCAGCAATGTCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTCAGAG 5520  
5869 CTACACCATCACTGGCTTACAAACAGGCACTGACTACAGATCTACCTGTACACCTTGAA 5928

5521 CTACACCATCACTGTTTACAAACAGGCACTGACTACAGATCTACCTGTACACCTTGAA 5580  
5929 TGCAATGCTGGAGCTCCCTGTGGTCACTGAGCCTTCACTGCCATTTGATGACCAATC 5988  
5581 TGCAATGCTGGAGCTCCCTGTGGTCACTGAGCCTTCACTGCCATTTGATGACCAATC 5640  
5989 CAACCTGGTTCCTGGCCACACACCCAAATTCCTTGTGTATCATGCGAGCGCCACG 6048  
5641 CAACCTGGTTCCTGGCCACACACCCAAATTCCTTGTGTATCATGCGAGCGCCACG 5700  
6049 TCCAGGATTAACCGCTACATCATCAAGTATGAGAAGCTGGTCTCTCCAGAGAAAT 6108  
5701 TCCAGGATTAACCGCTACATCATCAAGTATGAGAAGCTGGTCTCTCCAGAGAAAT 5760  
6109 GGTCCCTGGCCCGCCCTGTGTGTCACAGAGCTACTATTACTGGCCCTGGAACCGGGAAC 6168  
5761 GGTCCCTGGCCCGCCCTGTGTGTCACAGAGCTACTATTACTGGCCCTGGAACCGGGAAC 5820  
6169 CGAATATACAAATTTATGTCATTGGCCCTGAAGAAATTAATCAGAAGAGCGAGCCCTGATTGG 6228  
5821 CGAATATACAAATTTATGTCATTGGCCCTGAAGAAATTAATCAGAAGAGCGAGCCCTGATTGG 5880  
6229 AAGGAAAGAGCAGAGCTTCCCAACTGTGTAACTTCCACACCCCAATCTTCATGG 6288  
5881 AAGGAAAGAGCAGAGCTTCCCAACTGTGTAACTTCCACACCCCAATCTTCATGG 5940  
6289 ACCAGAGATCTTGAATGTTTCCCTTCCACAGTTTCAAAGACCCCTTTGTCACCCACCCCTGG 6348  
5941 ACCAGAGATCTTGAATGTTTCCCTTCCACAGTTTCAAAGACCCCTTTGTCACCCACCCCTGG 6000  
6349 GTATGACACTGGAATGTTTACAGTCTTCCGCACTTCTGTCAGCAACCCAGTGTGG 6408  
6001 GTATGACACTGGAATGTTTACAGTCTTCCGCACTTCTGTCAGCAACCCAGTGTGG 6060  
6409 GCAACAAATGATCTTTGAGGAACATGTTTAGCGGACACACCGCCCAACAAAGGCCAC 6468  
6061 GCAACAAATGATCTTTGAGGAACATGTTTAGCGGACACACCGCCCAACAAAGGCCAC 6120  
6469 CCCCATAAGGCATAGGCCAAGACCATATCCCGCGGATGTAGGAACAAGCTCTCTCTCA 6528  
6121 CCCCATAAGGCATAGGCCAAGACCATATCCCGCGGATGTAGGAACAAGCTCTCTCTCA 6180  
6529 GACAAACCATCTCATGGGCCCATTCAGGACACTTCTGAGTACATCTTTCATGTCTATCC 6588  
6181 GACAAACCATCTCATGGGCCCATTCAGGACACTTCTGAGTACATCTTTCATGTCTATCC 6240  
6589 TGTGGCACTGATGAAGAACCTTTACAGTTCAGGTTCTCGAACTTCTACAGTGCAC 6648  
6241 TGTGGCACTGATGAAGAACCTTTACAGTTCAGGTTCTCGAACTTCTACAGTGCAC 6300  
6649 TGTGACGCTCAGAGAGTGCACCATCAACATCATAGTGGAGGACCTGAAAGACCA 6708  
6301 TGTGACGCTCAGAGAGTGCACCATCAACATCATAGTGGAGGACCTGAAAGACCA 6360  
6709 GCAGAGGCATTAAGGTTCCGGAAGAGGTTGTTACCGTGGGCAACTCTGTCAACAGGCTT 6768  
6361 GCAGAGGCATTAAGGTTCCGGAAGAGGTTGTTACCGTGGGCAACTCTGTCAACAGGCTT 6420  
6769 GAAACCACTACGAGTGTCTGCTTGAACCTTACAGTGTCCCATTTATGCGGTGG 6828  
6421 GAAACCACTACGAGTGTCTGCTTGAACCTTACAGTGTCCCATTTATGCGGTGG 6480  
6829 AGATGAGTGGGAACCAATGCTGTAATCAGGCTTTAAACCTGTTGCGGAGTCTTTAGGCTT 6888  
6481 AGATGAGTGGGAACCAATGCTGTAATCAGGCTTTAAACCTGTTGCGGAGTCTTTAGGCTT 6540  
6889 TGAAGTGTCTATTTACAGATGATTCATCTAGATGCTGTCATGCAATGCTGTGAACCTA 6948  
6541 TGAAGTGTCTATTTACAGATGATTCATCTAGATGCTGTCATGCAATGCTGTGAACCTA 6600  
6949 CAAGATTTGAGAGAAAGTGGACCGCTCAGGAGAAATGGCCAGATGATGAGTGCACATG 7008



```

6601 CAAGATTGGAGAGAAAGTGGGACGTCAGGGAGAAATGCCAGATGATGAGCTGCACATG 6660
7009 TCTTGGGAACGGAAGGAGAAATTCAGTGTGACCTCATGAGGCAACGCTGTACGATGA 7068
6661 TCTTGGGAACGGAAGGAGAAATTCAGTGTGACCTCATGAGGCAACGCTGTACGATGA 6720
7069 TGGGAAGACATACCAGTAGGAGAAAGTGGGAGGAATATCTCGGTGCCATTTGCTC 7128
6721 TGGGAAGACATACCAGTAGGAGAAAGTGGGAGGAATATCTCGGTGCCATTTGCTC 6780
7129 CTGCATGCTTTGGAGGCGCAGGGGCTGGGCTGTGACAACTGCGCAGACCTGGGG 7188
6781 CTGCATGCTTTGGAGGCGCAGGGGCTGGGCTGTGACAACTGCGCAGACCTGGGG 6840
7189 TGAACCCAGTCCGGAAGGCACTACTGCGCAGTCTTACAAACGAGTATCTCAGAGATACA 7248
6841 TGAACCCAGTCCGGAAGGCACTACTGCGCAGTCTTACAAACGAGTATCTCAGAGATACA 6900
7249 TCAGAGAAACAAACACTAATGTTAATGGCCCAATGAGTGTCTTACGCTTTAGATGTACA 7308
6901 TCAGAGAAACAAACACTAATGTTAATGGCCCAATGAGTGTCTTACGCTTTAGATGTACA 6960
7309 GGCTGACAGAGAAAGATCCGAGAGTAATATCTTTCCAAATCCAGAGGAAACAGCATGT 7368
6961 GGCTGACAGAGAAAGATCCGAGAGTAATATCTTTCCAAATCCAGAGGAAACAGCATGT 7020
7369 CTCTGCGCAAGATCCATCTAAACTGGAGTGATGTTAGCAGACCCAGCTTAGAGTTCTTC 7428
7021 CTCTGCGCAAGATCCATCTAAACTGGAGTGATGTTAGCAGACCCAGCTTAGAGTTCTTC 7080
7429 TTTCTTTCTTAAGCCCTTTGCTCTGGAGGAAGTTCTCCAGCTTCACTCAACTCAGCT 7140
7489 TCTCCAGGATCACTCCGAGGTTCTGAGGTTTCTCAATAAGAGGCTGCACATT 7548
7141 TCTCCAGGATCACTCCGAGGTTTCTGAGGTTTCTCAATAAGAGGCTGCACATT 7200
7549 GCCTGTCTGCTTCGAGATATCAATACCGCTCAGTATTTTAAATGAAGTATCTTAAGA 7608
7201 GCCTGTCTGCTTCGAGATATCAATACCGCTCAGTATTTTAAATGAAGTATCTTAAGA 7260
7609 TTTGGTTGGGATCAATAGGAGAGCATATGCGCCCAACCAAGATCAATGTTTGAAT 7668
7261 TTTGGTTGGGATCAATAGGAGAGCATATGCGCCCAACCAAGATCAATGTTTGAAT 7320
7669 GATATGACCAAAATTTTAAAGTAGGAAAGTCACCCAAACACTTCTGCTTTCACTTAAGTCT 7728
7321 GATATGACCAAAATTTTAAAGTAGGAAAGTCACCCAAACACTTCTGCTTTCACTTAAGTCT 7380
7729 CTGCGCGGATCACTGTAGGAAACAGCATGATCTTGTACTGTGATATTTTAAATATCCA 7788
7381 CTGCGCGGATCACTGTAGGAAACAGCATGATCTTGTACTGTGATATTTTAAATATCCA 7440
7789 CAGTACT 7795
7441 CAGTACT 7447

RESULT 8
JS-10-182-936A-75
; Sequence 75, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A

```

```

; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-182-936A-75

Query Match 95.2%; Score 7423; DB 12; Length 7680;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 349 GAAGAGCAAGAGGCGAGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTGGCTGTCAGTCA 408
Db 1 GAAGAGCAAGAGGCGAGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTGGCTGTCAGTCA 60
QY 409 AAGCAAGCCCGGTTGTTATGACAAATGGAAAAACATATCAGATAAATCAACAGTGGGAGCG 468
Db 61 AAGCAAGCCCGGTTGTTATGACAAATGGAAAAACATATCAGATAAATCAACAGTGGGAGCG 120
QY 469 GACCTACTGAGCAATCGTGTGTTTGTCTTGTATGAGGAGAACCGAGGTTTAACTG 528
Db 121 GACCTACTGAGCAATCGTGTGTTTGTCTTGTATGAGGAGAACCGAGGTTTAACTG 180
QY 529 CGAAAGTAAACCTGAAGCTGAAGAGACTTCTTTGACAACTACATGAGGAACTTACCG 588
Db 181 CGAAAGTAAACCTGAAGCTGAAGAGACTTCTTTGACAACTACATGAGGAACTTACCG 240
QY 589 AGTGGGTGACACTTATGAGGCTCTTAAGACTCATGATCTGGAGCTGTACCTGCACTCGG 648
Db 241 AGTGGGTGACACTTATGAGGCTCTTAAGACTCATGATCTGGAGCTGTACCTGCACTCGG 300
QY 649 GGCTGGGCGAGGAGAAATAAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGTCACTC 708
Db 301 GGCTGGGCGAGGAGAAATAAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGTCACTC 360
QY 709 CTACAAGATTGGTGACACCTTGAGGAGACCACTGAGACTGCTGGTGTACATGTTAGAGTG 768
Db 361 CTACAAGATTGGTGACACCTTGAGGAGACCACTGAGACTGCTGGTGTACATGTTAGAGTG 420
QY 769 TGTGTGCTCTGTTAATGAAAAAGGAGAAATGGACCTGCAAGCCCATAGCTGAGAGAGTGT 828
Db 421 TGTGTGCTCTGTTAATGAAAAAGGAGAAATGGACCTGCAAGCCCATAGCTGAGAGAGTGT 480
QY 829 TGATCATGCTCTGGGACTTCTTATGTTGGTGAGAGAAACGTTGGAGAGAGCCCTACCAAGG 888
Db 481 TGATCATGCTCTGGGACTTCTTATGTTGGTGAGAGAAACGTTGGAGAGAGCCCTACCAAGG 540
QY 889 CTGGATGATGTTAGATTGTACTTGCCTGGGAGAGGAGGAGCGGACGCTACCTTGCATCTC 948
Db 541 CTGGATGATGTTAGATTGTACTTGCCTGGGAGAGGAGGAGCGGACGCTACCTTGCATCTC 600
QY 949 TAGAAATAGATGCAACGATCAGGACCAAGAGACATCTTATAGAAATGAGAGACCTGGAG 1008
Db 601 TAGAAATAGATGCAACGATCAGGACCAAGAGACATCTTATAGAAATGAGAGACCTGGAG 660
QY 1009 CAAGAAGGATTAATCGAGGAAACCTGCTCCAGTGTCTGACAGGCAACCGCGGAGAGA 1068
Db 661 CAAGAAGGATTAATCGAGGAAACCTGCTCCAGTGTCTGACAGGCAACCGCGGAGAGA 720
QY 1069 GTGGAAGTGTGAGAGGCGACACTCTGTGTCAGACCAATCGAGGAGATCTGGCCCTTCAC 1128
Db 721 GTGGAAGTGTGAGAGGCGACACTCTGTGTCAGACCAATCGAGGAGATCTGGCCCTTCAC 780

```



1129 CGATGTTCTGTCAGCTGTTTACCAACCGCAGCCTCACCCACGCTCCTCCCTATGGCCA 1188  
 781 CGATGTTCTGTCAGCTGTTTACCAACCGCAGCCTCACCCACGCTCCTCCCTATGGCCA 940  
 1189 CTGTGTCAAGACAGAGTGTGTGTCTACTCTGTGGGATGACGTGCTGAAGACACAGG 1248  
 841 CTGTGTCAAGACAGAGTGTGTGTCTACTCTGTGGGATGACGTGCTGAAGACACAGG 900  
 1249 AATTAAGCAATGCTTTGCAAGCTGCTGGGCAAGGAGTCAGTGCACAGAGACAGCTGT 1308  
 901 AATTAAGCAATGCTTTGCAAGCTGCTGGGCAAGGAGTCAGTGCACAGAGACAGCTGT 960  
 1309 AACCAGACTTACGCTGGCACTCAAAATGGAGAGCCATGTCTTACCAATTCACCTACAA 1368  
 961 AACCAGACTTACGCTGGCACTTAAATGGAGAGCCATGTCTTACCAATTCACCTACAA 1020  
 1369 TGGCAGAGCTTCTACTCTGCAAGCTGCTGGGCAAGGAGTCAGTGCACAGAGACAGCTGT 1428  
 1021 TGGCAGAGCTTCTACTCTGCAAGCTGCTGGGCAAGGAGTCAGTGCACAGAGACAGCTGT 1080  
 1429 CACAACCTTCAAAATTAAGCAGGACCAAGAAATACTCTTCTGACAGACCACTGTGTTT 1488  
 1081 CACAACCTTCAAAATTAAGCAGGACCAAGAAATACTCTTCTGACAGACCACTGTGTTT 1140  
 1489 GGTTCAGACTGAGGAGGAAATTCCAATGTGTCTGTGCACTTCCCTTCCATACAA 1548  
 1141 GGTTCAGACTGAGGAGGAAATTCCAATGTGTCTGTGCACTTCCCTTCCATACAA 1200  
 1549 CAACCAATTTACACTGATGCACTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1608  
 1201 CAACCAATTTACACTGATGCACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
 1609 GACCAACAGAACTATGATGCGGACCAAGAAATTTGGGTTCTGCGCCATGCTGCCACGA 1668  
 1261 GACCAACAGAACTATGATGCGGACCAAGAAATTTGGGTTCTGCGCCATGCTGCCACGA 1320  
 1669 GGAATCTGCAACCAATGAAGGGGTCAATGACGATTTGGAGATCAGTGGGATGAAGCA 1728  
 1321 GGAATCTGCAACCAATGAAGGGGTCAATGACGATTTGGAGATCAGTGGGATGAAGCA 1380  
 1729 GCATGATGGGTCAATGATGAGTGCAGCTGTTGGGAAATGGTGGTGGGAAATGGAC 1788  
 1381 GCATGATGGGTCAATGATGAGTGCAGCTGTTGGGAAATGGTGGTGGGAAATGGAC 1440  
 1789 ATGATGTCCTTACCTGAGCTTCCAGATCAGTGCATTTGATGATGATGATGATGATGATGAT 1848  
 1441 ATGATGTCCTTACCTGAGCTTCCAGATCAGTGCATTTGATGATGATGATGATGATGATGAT 1500  
 1849 GAACGACACATTTCCACAGGCTCATGAAGGGGACATGCTGAACTGTATGATGATGATGATGAT 1908  
 1501 GAACGACACATTTCCACAGGCTCATGAAGGGGACATGCTGAACTGTATGATGATGATGATGAT 1560  
 1909 TCAGGTTGGGGCAGGTGGAGTGTGATCCGTCGACCAATGCCAGATTCAGAGACTGG 1968  
 1561 TCAGGTTGGGGCAGGTGGAGTGTGATCCGTCGACCAATGCCAGATTCAGAGACTGG 1620  
 1969 GACGTTTATCAAAATGGAGATTCATGGGAGGATGATGATGATGATGATGATGATGATGATGAT 2028  
 1621 GACGTTTATCAAAATGGAGATTCATGGGAGGATGATGATGATGATGATGATGATGATGATGAT 1680  
 2029 CTACTGCTATGGCGTGGCATTGGGGAGTGGCATTGGCACTTACAGACCTTACAGACCTATCCAAG 2088  
 1681 CTACTGCTATGGCGTGGCATTGGGGAGTGGCATTGGCACTTACAGACCTTATCCAAG 1740  
 2089 CTCAGTGGTCTGCTGCAAGTATTTATCACTGAGACTCCGAGTCAGGCTCAGGCTCAGGCTCAGGCT 2148  
 1741 CTCAGTGGTCTGCTGCAAGTATTTATCACTGAGACTCCGAGTCAGGCTCAGGCTCAGGCTCAGGCT 1800  
 2149 CATCAGTGGAAATGACACAGGCACTTCAATTTCCAAATGATGATGATGATGATGATGATGATGAT 2208  
 1801 CATCAGTGGAAATGACACAGGCACTTCAATTTCCAAATGATGATGATGATGATGATGATGATGAT 1860  
 2209 TAAAAATTTCTGTAGGCCCTTGGAGGAGGCTACCATACCGAGGCTTAAACTCCTACAC 2268

1861 TAAAAATTTCTGTAGGCCCTTGGAGGAGGCTACCATACCGAGGCTTAAACTCCTACAC 1920  
 2269 CATCAAGGCTGAAGCTTGGTGTATGAGGGGCGAGCTATCAGCATCCAGCAGTA 2328  
 1921 CATCAAGGCTGAAGCTTGGTGTATGAGGGGCGAGCTATCAGCATCCAGCAGTA 1980  
 2329 CGGCCACCAAGAGTGAAGTCTGCTTGTGCTTCCACCAACAGGACCCAGCAGCAGCTGTGAC 2388  
 1981 CGGCCACCAAGAGTGAAGTCTGCTTGTGCTTCCACCAACAGGACCCAGCAGCAGCTGTGAC 2040  
 2389 CAGCAACCCGTGACAGAGAGAGCTCCCTTTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2448  
 2041 CAGCAACCCGTGACAGAGAGAGCTCCCTTTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
 2449 TGTGACCCCAAAATCAAGCCAGTGTGCTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2508  
 2101 TGTGACCCCAAAATCAAGCCAGTGTGCTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
 2509 GTGGGATTCGGGTGGAATATGAGCTGAGTGAGGAGGAGATGAGCCAGCAGTACCTGGA 2568  
 2161 GTGGGATTCGGGTGGAATATGAGCTGAGTGAGGAGGAGATGAGCCAGCAGTACCTGGA 2220  
 2569 TCTTCCAAAGCACAGCCACTTCTGTGAACTCCCTGACCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2628  
 2221 TCTTCCAAAGCACAGCCACTTCTGTGAACTCCCTGACCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
 2629 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTTTGTATCTGCTGCTGCTGCTGCTGCTGCT 2688  
 2281 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTTTGTATCTGCTGCTGCTGCTGCTGCTGCT 2340  
 2689 AACACAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2748  
 2341 AACACAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400  
 2749 TGTGTTTCTGCTGAGCAGACCCAGGCTCCCATCAGAGGTACAGAAATGATCTATTCGCC 2808  
 2401 TGTGTTTCTGCTGAGCAGACCCAGGCTCCCATCAGAGGTACAGAAATGATCTATTCGCC 2460  
 2809 ATCAGTAGAGGTAGCAGCAGAGACTCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2868  
 2461 ATCAGTAGAGGTAGCAGCAGAGACTCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
 2869 CAGTGACTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2928  
 2521 CAGTGACTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580  
 2929 AGAAAGTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2988  
 2581 AGAAAGTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640  
 2989 GCCCTCTCCAGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3048  
 2641 GCCCTCTCCAGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
 3049 GACACCGCTGAGAGTGACGTGACCGGTACCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3108  
 2701 GACACCGCTGAGAGTGACGTGACCGGTACCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760  
 3109 TGGCAGCAGCAGGAGAGGCTGCCATCAGCAGGAGAACCTTTGAGAGGTCAGGAGTCAACCGGCT 3168  
 2761 TGGCAGCAGCAGGAGAGGCTGCCATCAGCAGGAGAACCTTTGAGAGGTCAGGAGTCAACCGGCT 2820  
 3169 GTCCCTGAGGTCACCTATTAATCTCAAAAGTCTTTGAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 3228  
 2821 GTCCCTGAGGTCACCTATTAATCTCAAAAGTCTTTGAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2880  
 3229 GCCTGAGTGTCAACAGACCAACTGAGTCTCCCACTAACCTTCCAGTCTTGTCAA 3288  
 2881 GCCTGAGTGTCAACAGACCAACTGAGTCTCCCACTAACCTTCCAGTCTTGTCAA 2940  
 3289 TGAATCTGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3348

2941 TGAACGTGATTCTACTGTCTCTGGTGAAGTGGATCCACCTCGGGCCCGAGATAACAGGATA 3000  
3349 CCGACTGACCGTGGGCTTTACCCGAGAGAGACAGCCAGCGAGTACAATGTGGTCCCTC 3408  
3001 CCGACTGACCGTGGGCTTTACCCGAGAGAGAGCCAGCCAGCGAGTACATGTGGTCCCTC 3060  
3409 TGTCTCCAGTACCCACTGAGGAATCTGAGCGCTGTCATCTGAGTACACCGTATCCCTGT 3468  
3061 TGTCTCCAGTACCCCTGAGGAATCTGAGCGCTGTCATCTGAGTACACCGTATCCCTGT 3120  
3469 GGCCTATAGGGCAACCAAGAGAGCCCAAGGCACTGGAGTCTTTACACACTGCGAGCC 3528  
3121 GGCCTATAGGGCAACCAAGAGAGCCCAAGGCACTGGAGTCTTTACACACTGCGAGCC 3180  
3529 TGGGAGCTCTATTCCACTTTACACACCGAGTGTACTGAGACCAACCATTTGATCACATG 3588  
3181 TGGGAGCTCTATTCCACTTTACACACCGAGTGTACTGAGACCAACCATTTGATCACATG 3240  
3589 GAGCGCTGCTCCAGGAATGGTTTAAAGCTGGGTGTACGACCAAGCCAGGAGGAGGC 3648  
3241 GAGCGCTGCTCCAGGAATGGTTTAAAGCTGGGTGTACGACCAAGCCAGGAGGAGGC 3300  
3649 ACCACGAGAGTGAATTCAGACTCAGGAAGCATCGTTGTCTCGGTTGACTCCAGGAGT 3708  
3301 ACCACGAGAGTGAATTCAGACTCAGGAAGCATCGTTGTCTCGGTTGACTCCAGGAGT 3360  
3709 AGAATAGCTCTACACCATCCAGTCTGAGAGTGGACAGGAAGAGATGCCCAATGT 3768  
3361 AGAATAGCTCTACACCATCCAGTCTGAGAGTGGACAGGAAGAGATGCCCAATGT 3420  
3769 AAACAAAGTGGTGACACCATTTGTCTCCACCAACAACTTGCATCTGAGGCAACCTGA 3828  
3421 AAACAAAGTGGTGACACCATTTGTCTCCACCAACAACTTGCATCTGAGGCAACCTGA 3480  
3829 CACTGGAGTCTCAGTCTCTGGAGAGGAGACCAACCCAGACATTACTGTTATAG 3888  
3481 CACTGGAGTCTCAGTCTCTGGAGAGGAGACCAACCCAGACATTACTGTTATAG 3540  
3889 AATTACCAACCCCTACAAACCGCCAGCAGGGAATTTCTTGGAGAAAGTGTCTCATGC 3948  
3541 AATTACCAACCCCTACAAACCGCCAGCAGGGAATTTCTTGGAGAAAGTGTCTCATGC 3600  
3949 TGATCAGAGTCTCTGACATTTGATACCTGAGTCCGGCTCGAGTACAAATGTAGTGT 4008  
3601 TGATCAGAGTCTCTGACATTTGATACCTGAGTCCGGCTCGAGTACAAATGTAGTGT 3660  
4009 TTACACTGTCAAGATGACAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 4068  
3661 TTACACTGTCAAGATGACAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 3720  
4069 TCCTCTCCCACTGACCTGCGATTACCAACATTTGATGAGTGTCCAGACCACTGTCACCTG 4128  
3721 TCCTCTCCCACTGACCTGCGATTACCAACATTTGATGAGTGTCCAGACCACTGTCACCTG 3780  
4129 GGTCTCAACCCCTCCATTCATTAACCAACTTCTGTTGCGTTACTCATCTGTGAAAAA 4188  
3781 GGTCTCAACCCCTCCATTCATTAACCAACTTCTGTTGCGTTACTCATCTGTGAAAAA 3840  
4189 TGAGGAAGATGTTGACAGTGTCAATTTCTCTTCAAGAACATGAGTGTCTTAACAAA 4248  
3841 TGAGGAAGATGTTGACAGTGTCAATTTCTCTTCAAGAACATGAGTGTCTTAACAAA 3900  
4249 TCCTCTGCTGTGATCAGAAATGATGAGTGTCTCCAGTGTCTACGAAACATCAGAG 4308  
3901 TCCTCTGCTGTGATCAGAAATGATGAGTGTCTCCAGTGTCTACGAAACATCAGAG 3960  
4309 CACAACCTTTAGAGGAAGACAGAAAAAGTCTTGATTTCCCAACTGGCAATGACATTTTC 4368  
3961 CACAACCTTTAGAGGAAGACAGAAAAAGTCTTGATTTCCCAACTGGCAATGACATTTTC 4020  
4369 TGATATTACTGCCAATCTTTTACTGTGCACTGGATTGCTCTCGAGGCCACCATCACTG 4428  
4021 TGATATTACTGCCAATCTTTTACTGTGCACTGGATTGCTCTCGAGGCCACCATCACTG 4080

4429 CTACAGGATCGCCATCATCCCGAGCACTTCACTGGAGACCTCGAGAGATCGGTGCC 4488  
4081 CTACAGGATCGCCATCATCCCGAGCACTTCACTGGAGACCTCGAGAGATCGGTGCC 4140  
4489 CCACCTCTCGGAATTCATCACCTCTCACACCTCTCTCCAGGACACAGATGATGTGTGAG 4548  
4141 CCACCTCTCGGAATTCATCACCTCTCACACCTCTCTCCAGGACACAGATGATGTGTGAG 4200  
4549 CATCGTTGCTCTTAATGGCAGAGAGAAAGTCCCTTATTGATTGGCCAAACAATCAACAGT 4608  
4201 CATCGTTGCTCTTAATGGCAGAGAGAAAGTCCCTTATTGATTGGCCAAACAATCAACAGT 4260  
4609 TTCTGATGTTCCGAGGACCTGGAGTGTGTGGACCCGCCACGACCTACTGATCAG 4668  
4261 TTCTGATGTTCCGAGGACCTGGAGTGTGTGGACCCGCCACGACCTACTGATCAG 4320  
4669 CTGGGATGCTCTGCTCTCTCACAGTGTGATATTACAGATCACTTACGGAGAAAACAGAGG 4728  
4321 CTGGGATGCTCTGCTCTCTCACAGTGTGATATTACAGATCACTTACGGAGAAAACAGAGG 4380  
4729 AANTAGCCCTGTCCAGAGTTCACGTCTGGGAGCAAGTCTACAGTACCATCAGCGG 4788  
4381 AANTAGCCCTGTCCAGAGTTCACGTCTGGGAGCAAGTCTACAGTACCATCAGCGG 4440  
4789 CCTTAAACCTGGAGTTCATTAACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAG 4848  
4441 CCTTAAACCTGGAGTTCATTAACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAG 4500  
4849 CCGCGAAGCAGCAGCAATTTCCATTAATACCGAACAAGAAATTCACAAAACATCCCA 4908  
4501 CCGCGAAGCAGCAGCAATTTCCATTAATACCGAACAAGAAATTCACAAAACATCCCA 4560  
4909 GATGCAAGTGAACCATGTTTCAAGACAAACAGATTAAGTGTCAAGTGGCTGCTTCAAGTTC 4968  
4561 GATGCAAGTGAACCATGTTTCAAGACAAACAGATTAAGTGTCAAGTGGCTGCTTCAAGTTC 4620  
4969 CCCTGTTACTGGTTTACAGAGTAACCACTCCCAAAAATGGACAGGACCAACAAAAC 5028  
4621 CCCTGTTACTGGTTTACAGAGTAACCACTCCCAAAAATGGACAGGACCAACAAAAC 4680  
5029 TAAACCTGACGCTCCAGATCAACACAGAAATGATATTGAAAGGCTTGACGCCACAGTGA 5088  
4681 TAAACCTGACGCTCCAGATCAACACAGAAATGATATTGAAAGGCTTGACGCCACAGTGA 4740  
5089 GTATGTGGTTAGTGTCTATGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGTTTCAAG 5148  
4741 GTATGTGGTTAGTGTCTATGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGTTTCAAG 4800  
5149 TGCAGTAACCAACATTCATCGCTTAAAGACTGGCAATTCATGATGTGATGTGCGATTTC 5208  
4801 TGCAGTAACCAACATTCATCGCTTAAAGACTGGCAATTCATGATGTGATGTGCGATTTC 4860  
5209 CATCAAAATTTGCTTTGGGAAAGCCCAAGGGGCAAGTTTCCAGGTFACAGGTFACCTACTC 5268  
4861 CATCAAAATTTGCTTTGGGAAAGCCCAAGGGGCAAGTTTCCAGGTFACAGGTFACCTACTC 4920  
5269 GAGCCCTGAGAGATGGAATCCATGAGCTATTCCCTGACCTGATGGTGAAGNAGACATGC 5328  
4921 GAGCCCTGAGAGATGGAATCCATGAGCTATTCCCTGACCTGATGGTGAAGNAGACATGC 4980  
5329 AGAGCTGCAAGGCTCTCAGACCGGGTCTTGAGTACACAGTCAAGTGTGTTGCTTTCACGA 5388  
4981 AGAGCTGCAAGGCTCTCAGACCGGGTCTTGAGTACACAGTCAAGTGTGTTGCTTTCACGA 5040  
5389 TGATATGAGAGACCGAGCCCTGATTGGAAACCCAGTCCACAGCTATTTCCTGCACCACTGA 5448  
5041 TGATATGAGAGACCGAGCCCTGATTGGAAACCCAGTCCACAGCTATTTCCTGCACCACTGA 5100  
5449 CCTGAAGTTCCTCAGGTTCACACCCCAAGCCTGAGGCGCCAGTGGACACCAACCAATGT 5508  
5101 CCTGAAGTTCCTCAGGTTCACACCCCAAGCCTGAGGCGCCAGTGGACACCAACCAATGT 5160

QY 5509 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5568  
DB 5161 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5220  
QY 5569 AATCAACCTTCTCTCTGACAGCTCATCCGTGGTGTATCAGGACTTATGGTGGCCACCAA 5628  
DB 5221 AATCAACCTTCTCTCTGACAGCTCATCCGTGGTGTATCAGGACTTATGGTGGCCACCAA 5280  
QY 5629 ATATGAGTGTAGTGTATGCTCTTAAGGACACTTTGACAGCAGACGACTCAGGAGT 5688  
DB 5281 ATATGAGTGTAGTGTATGCTCTTAAGGACACTTTGACAGCAGACGACTCAGGAGT 5340  
QY 5689 TGTCAACCACTCTGGAGAAATGTGAGCCCAACCAAGAGGCTCGTGTGACAGATGCTACTGA 5748  
DB 5341 TGTCAACCACTCTGGAGAAATGTGAGCCCAACCAAGAGGCTCGTGTGACAGATGCTACTGA 5400  
QY 5749 GACACCATATCACTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTCCAGTTGA 5808  
DB 5401 GACCACCATATCACTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTCCAGTTGA 5460  
QY 5809 TGCCGTTCCAGCAATGAGCCAGACTCCCAATCCAGAGAACCAATCAAGCCAGATGTCAGAAG 5868  
DB 5461 TGCCGTTCCAGCAATGAGCCAGACTCCCAATCCAGAGAACCAATCAAGCCAGATGTCAGAAG 5520  
QY 5869 CTACACCATATCTGCTTTACACAGGCACTGACTCAAGATCTACCTGTACACCTTGAA 5928  
DB 5521 CTACACCATATCAAGTTTACAAACAGGCACTGACTCAAGATCTACCTGTACACCTTGAA 5580  
QY 5929 TGACAACTCTGGAGCTCCCTGTGTGTCATCAGCGCTCCACTGCCATTTGATGACCACTC 5988  
DB 5581 TGACAACTCTGGAGCTCCCTGTGTGTCATCAGCGCTCCACTGCCATTTGATGACCACTC 5640  
QY 5989 CAACCTGGGTTTCTTGGCCACCAACCACTTCTTGTCTGTATCATGGCAGCGCCCAAG 6048  
DB 5641 CAACCTGGGTTTCTTGGCCACCAACCACTTCTTGTCTGTATCATGGCAGCGCCCAAG 5700  
QY 6049 TGCCAGGATTAACCGCTACATCATCAAGTATGAGAGCGCTGGTCTCTCTCCACAGAGT 6108  
DB 5701 TGCCAGGATTAACCGCTACATCATCAAGTATGAGAGCGCTGGTCTCTCTCCACAGAGT 5760  
QY 6109 GGTCCCTCGGCCCCCTGTGTGTGTCACAGAGGCTACTATTACTGGCCTGGAAACCGGGAAC 6168  
DB 5761 GGTCCCTCGGCCCCCTGTGTGTGTCACAGAGGCTACTATTACTGGCCTGGAAACCGGGAAC 5820  
QY 6169 CGAATATACAAATTTATGTCTATGCTCCCTCAAGAAATATCAGAGAGCGAGCCCTGATTGG 6228  
DB 5821 CGAATATACAAATTTATGTCTATGCTCCCTCAAGAAATATCAGAGAGCGAGCCCTGATTGG 5880  
QY 6229 AAGGAAAAGACAGACGAGCTTCCCAACTGGTAACCTTCCACACCCCAATCTTCTATGG 6288  
DB 5881 AAGGAAAAGACAGACGAGCTTCCCAACTGGTAACCTTCCACACCCCAATCTTCTATGG 5940  
QY 6289 ACCAGAGATCTTGGATGTTCTTCCACAGTTCAAAAGACCCCTTGTGTCACCCACCTGG 6348  
DB 5941 ACCAGAGATCTTGGATGTTCTTCCACAGTTCAAAAGACCCCTTGTGTCACCCACCTGG 6000  
QY 6349 GPATCAGACTGGAATGTTATGCTTCCCTGGCACTTCTGGTCAGCAACCCAGTGTGG 6408  
DB 6001 GPATCAGACTGGAATGTTATGCTTCCCTGGCACTTCTGGTCAGCAACCCAGTGTGG 6060  
QY 6409 GCAACAAATGATCTTTGAGGAACATGTTTATAGGCGGACCAACCGGCCCAACAAAGGCGCAC 6468  
DB 6061 GCAACAAATGATCTTTGAGGAACATGTTTATAGGCGGACCAACCGGCCCAACAAAGGCGCAC 6120  
QY 6469 CCCCATAAGGCATAGGCCAAGACCATACCCCGGAATGTAGGACAAAGAGCTCTCTCTCA 6528  
DB 6121 CCCCATAAGGCATAGGCCAAGACCATACCCCGGAATGTAGGACAAAGAGCTCTCTCTCA 6180  
QY 6529 GACAAACCATCTCATGGGCCCACTTCCAGGACACTTCTGAGTACATCATTTTATGTCATCC 6588  
DB 6181 GACAAACCATCTCATGGGCCCACTTCCAGGACACTTCTGAGTACATCATTTTATGTCATCC 6240  
QY 6589 TGTGGCACTGATGAGAACCCCTTACAGTTGAGGTTCTTGGAGCTTCTTACAGTGGCAC 6648

DB 6241 TGTGGCACTGATGAGAAACCTTACAGTTGAGGTTCTTGGAACTTCTTACAGTGGCAC 6300  
QY 6649 TCTGACAGGCTCTCACACAGAGTGGCCACCTACAAATCATAGTGGAGGCACTGAAAGACCA 6708  
DB 6301 TCTGACAGGCTCTCACACAGAGTGGCCACCTACAAATCATAGTGGAGGCACTGAAAGACCA 6360  
QY 6709 GCAGAGGCATAGGTTCCGGAGAGAGGTTGTACCGTGGGCAACTCTGTCAACGAGAGGCTT 6768  
DB 6361 GCAGAGGCATAGGTTCCGGAGAGAGGTTGTACCGTGGGCAACTCTGTCAACGAGAGGCTT 6420  
QY 6769 GAAACCAACCTACGAGTACTCGTGTGATGACCCCTACACAGTGTCCCAATATGCGCTTGG 6828  
DB 6421 GAAACCAACCTACGAGTACTCGTGTGATGACCCCTACACAGTGTCCCAATATGCGCTTGG 6480  
QY 6829 AGATGAGTGGGAAACCAATGCTCTGAATCAGGCTTTAAACTGTGTGTCGCCAGTGTAGGCTT 6888  
DB 6481 AGATGAGTGGGAAACCAATGCTCTGAATCAGGCTTTAAACTGTGTGTCGCCAGTGTAGGCTT 6540  
QY 6889 TGGAAATGCTCATTTTCAGATGTATCTAGATGGTGCATGACAAATGGTGTGAACTA 6948  
DB 6541 TGGAAATGCTCATTTTCAGATGTATCTAGATGGTGCATGACAAATGGTGTGAACTA 6600  
QY 6949 CAAGATTTGGAGAGAGTGGGACCGTCAAGGAGAAATGAGGAGAAATGATGAGTGCACATG 7008  
DB 6601 CAAGATTTGGAGAGAGTGGGACCGTCAAGGAGAAATGAGGAGAAATGATGAGTGCACATG 6660  
QY 7009 TCTTGGGAAACGGAAGAGAAATCAAGTGTGACCCCTCATGAGGCAACGTTTACGATGA 7068  
DB 6661 TCTTGGGAAACGGAAGAGAAATCAAGTGTGACCCCTCATGAGGCAACGTTTACGATGA 6720  
QY 7069 TGGGAAAGACATCACAGTAGGAGAACAGTGGCAGAGAAATATCTCGTGCATTTGCTC 7128  
DB 6721 TGGGAAAGACATCACAGTAGGAGAACAGTGGCAGAGAAATATCTCGTGCATTTGCTC 6780  
QY 7129 CTGCACTCTTTGGAGGCGAGCGGCTGTGACAACTGCCGAGACCTGGGG 7188  
DB 5781 CTGCACTCTTTGGAGGCGAGCGGCTGTGACAACTGCCGAGACCTGGGG 6840  
QY 7189 TGAAACCACTCCCGAAGGCACTACTGGCCAGTCTTCAACACAGTATTTCTCAGAGATACCA 7248  
DB 6841 TGAAACCACTCCCGAAGGCACTACTGGCCAGTCTTCAACACAGTATTTCTCAGAGATACCA 6900  
QY 7249 TCAGAGAAACAAACACTTAATGTTAAATGGCCCAATTCAGTGTCTTCAAGCCCTTTAGATGTACA 7308  
DB 5901 TCAGAGAAACAAACACTTAATGTTAAATGGCCCAATTCAGTGTCTTCAAGCCCTTTAGATGTACA 6960  
QY 7309 GGCTGACAGAGAGATTTCCCGAGAGTAAATCATCTTTCCAAATCCAGAGAAACAGCATGT 7368  
DB 6961 GGCTGACAGAGAGATTTCCCGAGAGTAAATCATCTTTCCAAATCCAGAGAAACAGCATGT 7020  
QY 7369 CTCTCTGCCAAGATCCATCTAACTGGAGTGTATGAGCAGACCCAGCTTAGAGTTCTTC 7428  
DB 7021 CTCTCTGCCAAGATCCATCTAACTGGAGTGTATGAGCAGACCCAGCTTAGAGTTCTTC 7080  
QY 7429 TTTCTTTTCTTAAAGCCCTTTGCTCTGGAGGAAGTTCTTCCAGCTTCAGCTCAACTCAGCT 7488  
DB 7081 TTTCTTTTCTTAAAGCCCTTTGCTCTGGAGGAAGTTCTTCCAGCTTCAGCTCAACTCAGCT 7140  
QY 7489 TCTCCAGCATCACCCCTGGAGTTTCTGAGGGTTTCTCATAAATGAGGGCTGCACATT 7548  
DB 7141 TCTCCAGCATCACCCCTGGAGTTTCTGAGGGTTTCTCATAAATGAGGGCTGCACATT 7200  
QY 7549 GCCTGTCTTCTGCTTCCGAAGTATTCAAATACCGCTCAGTATTTTAAATGAAGTATTTAAAGA 7608  
DB 7201 GCCTGTCTTCTGCTTCCGAAGTATTCAAATACCGCTCAGTATTTTAAATGAAGTATTTAAAGA 7260  
QY 7609 TTTGGTTTGGGATCAATAGGAAAGCATATGCGACCAACCAAGATGCAATGTTTGAAT 7668  
DB 7261 TTTGGTTTGGGATCAATAGGAAAGCATATGCGACCAACCAAGATGCAATGTTTGAAT 7320  
QY 7669 GATATGACCAAAATTTTAAATAGGAAAGTCAACCAACCACTTCTGCTTTCACTTAAAGTGT 7728







4009 TTACTGTCAAGGATGACAAAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 4068  
3661 TTACTGTCAAGGATGACAAAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 3720  
4069 TCCTCTCCCTGACCTGCGGATTCACCAACATTTGGTCCAGACACCATGCGTGTCACTG 4128  
3721 TCCTCTCCCTGACCTGCGGATTCACCAACATTTGGTCCAGACACCATGCGTGTCACTG 3780  
4129 GGCTCCACCCCATCCATGTATTTAACCAACTTCTGTGTGGTGTACTCACCTGTGA AAAA 4188  
3781 GGCTCCACCCCATCCATGTATTTAACCAACTTCTGTGTGGTGTACTCACCTGTGA AAAA 3840  
4189 TGAGAGAGATGTTGACAGATGTTCAATTTCTCTTCCATGACAGATGCGTGTCTTAA CAAA 4248  
3841 TGAGAGAGATGTTGACAGATGTTCAATTTCTCTTCCATGACAGATGCGTGTCTTAA CAAA 3900  
4249 TCTCTGCTGCTGACAGATATGTTAGTGAAGTGTCTCCAGTGTCTTACAGCAACATGAGAG 4308  
3901 TCTCTGCTGCTGACAGATATGTTAGTGAAGTGTCTCCAGTGTCTTACAGCAACATGAGAG 3960  
4309 CACACTCTTAGAGAGACAGAAACAGGTCTTGATTTCCCAACTGCGCATTTGACTTTTC 4368  
3961 CACACTCTTAGAGAGACAGAAACAGGTCTTGATTTCCCAACTGCGCATTTGACTTTTC 4020  
4369 TGATATTTACTGCAACTCTTTTACTGTGTGACACTGGATTTCTCTCGAGCCACCATCACTGG 4428  
4021 TGATATTTACTGCAACTCTTTTACTGTGTGACACTGGATTTCTCTCGAGCCACCATCACTGG 4080  
4429 CTACAGGATCCCGCATCATCCCGACACTTCTAGTGGAGACCTCGAGAGATCCGGTGCC 4488  
4081 CTACAGGATCCCGCATCATCCCGACACTTCTAGTGGAGACCTCGAGAGATCCGGTGCC 4140  
4489 CCACCTCTCGAAATTCATCACCTCTCACCACTCTACTCCAGSCACAGAGTATGTGTCAG 4548  
4141 CCACCTCTCGAAATTCATCACCTCTCACCACTCTACTCCAGSCACAGAGTATGTGTCAG 4200  
4549 CATGTTGTCTTATGAGAGAGAGAAAGTCCCTTATTTGATTTGGCCAACTCAACAGT 4608  
4201 CATGTTGTCTTATGAGAGAGAGAAAGTCCCTTATTTGATTTGGCCAACTCAACAGT 4260  
4609 TTCTGATGTTCCAGGGACCTGGAAAGTGTCTGCTGCGACCCCTCCAGCTACTGATCAG 4668  
4261 TTCTGATGTTCCAGGGACCTGGAAAGTGTCTGCTGCGACCCCTCCAGCTACTGATCAG 4320  
4669 CTGGGATGCTCTGCTGTGACAGTGAATATTTACAGATCATCTACGAGAAACAGGAGG 4728  
4321 CTGGGATGCTCTGCTGTGACAGTGAATATTTACAGATCATCTACGAGAAACAGGAGG 4380  
4729 AAATAGCCCTGTCCAGGAGTTCACCTGTGCTGGAGCAAGTCTACAGCTACCATCAGCGG 4788  
4381 AAATAGCCCTGTCCAGGAGTTCACCTGTGCTGGAGCAAGTCTACAGCTACCATCAGCGG 4440  
4789 CCTTAAACCTGGAGTTGATATACATCTACTGTGTATGCTGTCACTGGCGTGGAGACAG 4848  
4441 CCTTAAACCTGGAGTTGATATACATCTACTGTGTATGCTGTCACTGGCGCTGGAGACAG 4500  
4849 CCCCAGACAGCAAGCCCAATTTCCATTAATTTACCGAAGCAAGAAATTTGACAAACCATCCA 4908  
4501 CCCCAGACAGCAAGCCCAATTTCCATTAATTTACCGAAGCAAGAAATTTGACAAACCATCCA 4560  
4909 GATGCAAGTGAACCGATGTTTACGACAAACAGATTTAGTGTCAAGTGGCTGCTTCAAGTTTC 4968  
4561 GATGCAAGTGAACCGATGTTTACGACAAACAGATTTAGTGTCAAGTGGCTGCTTCAAGTTTC 4620  
4969 CCTGTTTACTGTTTACAGAGTACCACTCTCCCAAAATGGACAGGACCAACAAAAC 5028  
4621 CCTGTTTACTGTTTACAGAGTACCACTCTCCCAAAATGGACAGGACCAACAAAAC 4680  
5029 TAAAACTGCAAGTCCAGATCAACAGAAATGACTATTGAAGCTTTGACGCCCAACAGTGA 5088  
4681 TAAAACTGCAAGTCCAGATCAACAGAAATGACTATTGAAGCTTTGACGCCCAACAGTGA 4740  
5089 GTATGTGGTTAGTGTCTATGCTCAGAAATCCAGCGGAGAGAGTCAAGCTCTGGTTCAGAC 5148

4741 GTATGTGGTTAGTGTCTATGCTCAGAAATCCAGCGGAGAGAGTCAAGCTCTGGTTCAGAC 4800  
5149 TGCAAGTAAACCAATGATGCGCCCTAAAGGACTGGCATTTACTGATGTGGATGTGATTC 5208  
4801 TGCAAGTAAACCAATGATGCGCCCTAAAGGACTGGCATTTACTGATGTGGATGTGATTC 4860  
5209 CATCAAAATTTGTTGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTC 5268  
4861 CATCAAAATTTGTTGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTC 4920  
5269 GAGCCCTGAGGATGGAATCCATGAGCTATTCCTCGCACTGATGTGTAAGAGACACTGTC 5328  
4921 GAGCCCTGAGGATGGAATCCATGAGCTATTCCTCGCACTGATGTGTAAGAGACACTGTC 4980  
5329 AGAGCTGCAAGGCTCAGACCGGTTTCTGAGTACACAGTCAAGTGTGGTGTGCTTGCACGA 5388  
4981 AGAGCTGCAAGGCTCAGACCGGTTTCTGAGTACACAGTCAAGTGTGGTGTGCTTGCACGA 5040  
5389 TGATATGGAGAGCCAGCCCTGATTGGAAACCCAGTCCACAGCTATTTCTGTGACCAACTGA 5448  
5041 TGATATGGAGAGCCAGCCCTGATTGGAAACCCAGTCCACAGCTATTTCTGTGACCAACTGA 5100  
5449 CTTGAAGTTCACTCAGGTCACACCCACAGGCTGAGCGCCAGTGAGACACCACTCAATGT 5508  
5101 CTTGAAGTTCACTCAGGTCACACCCACAGGCTGAGCGCCAGTGAGACACCACTCAATGT 5160  
5509 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCACTCAATGAAAGA 5568  
5161 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCACTCAATGAAAGA 5220  
5569 AATCAACCTTGTCTGACAGCTCATCCGTTGTTGATCAGGACTTATGTTGGCCACCA 5628  
5221 AATCAACCTTGTCTGACAGCTCATCCGTTGTTGATCAGGACTTATGTTGGCCACCA 5280  
5629 ATATGAAGTGAAGTGTCTATGCTCTTAAGGACACTTTTGACAAGCAGACCACTGAGGAGT 5688  
5281 ATATGAAGTGAAGTGTCTATGCTCTTAAGGACACTTTTGACAAGCAGACCACTGAGGAGT 5340  
5689 TGTCACCACTCTGGAGATGTGAGCCCAACAGAGGGCTGCTGTGACAGATGCTACTGA 5748  
5341 TGTCAACCACTCTGGAGATGTGAGCCCAACAGAGGGCTGCTGTGACAGATGCTACTGA 5400  
5749 GACCACCATCACCATTAGCTGGAGAACCAAGACTGAGAGCACTCACTGGGCTCCAACTTGA 5808  
5401 GACCACCATCACCATTAGCTGGAGAACCAAGACTGAGAGCACTCACTGGGCTCCAACTTGA 5460  
5809 TGCCGTTTCCAGCAATGCGGCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAG 5868  
5461 TGCCGTTTCCAGCAATGCGGCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAG 5520  
5869 CTACACCATCAGCTGGCTTACCAACAGGCACTGACTACAGATCTACTGTACACCTTGAA 5928  
5521 CTACACCATCAGCTGGCTTACCAACAGGCACTGACTACAGATCTACTGTACACCTTGAA 5580  
5929 TGACAATGCTCGAGGCTCCCTGTGTCATGCAAGCTCCATGCGCATTTGATGCACTC 5988  
5581 TGACAATGCTCGAGGCTCCCTGTGTCATGCAAGCTCCATGCGCATTTGATGCACTC 5640  
5989 CAACCTCGGTTTCTGGCCACCAACCCCAATTCCTGTGCTGATCATGCGAGCGCCACG 6048  
5641 CAACCTCGGTTTCTGGCCACCAACCCCAATTCCTGTGCTGATCATGCGAGCGCCACG 5700  
6049 TGCCAGGATTAACGGCTTACATCATCAAGTATGAGAGCTGGGTCTCTCTCCAGAGAGT 6108  
5701 TGCCAGGATTAACGGCTTACATCATCAAGTATGAGAGCTGGGTCTCTCTCCAGAGAGT 5760  
6109 GGTCCCTCGGCCCGCCCTGGTGTCTCAGAGGCTACTATTACTGGCTGTGAACCGGGAAC 6168  
5761 GGTCCCTCGGCCCGCCCTGGTGTCTCAGAGGCTACTATTACTGGCTGTGAACCGGGAAC 5820  
6169 CGAATATACAAATTTATGCTCATTTGCTGCTGAAAGAAATATCAGAGAGCGAGCCCTGATTGG 6228



5821	CGAATATACAA	TTTATGTCTAT	TCCTCCCTGAAGAATAT	ATACGAAAGCGAGCCCTGATTGG	5888
6229	AAGGAAAAAGAC	AGACGAGCTT	CCCCAACTGGTAACCTT	CCACACCCCAATCTTCATGG	6288
5881	AAGGAAAAAGAC	AGACGAGCTT	CCCCAACTGGTAACCTT	CCACACCCCAATCTTCATGG	5940
6289	ACCAGAGATCT	TGGATGTTCT	CTTCCACAGTTC	CAAAAGACCCCTTTCGTACCCACACCTGG	6348
5941	ACACAGATCT	TGGATGTTCT	CTTCCACAGTTC	CAAAAGACCCCTTTCGTACCCACACCTGG	6000
6349	GTATGACATCG	TGAAAAATGG	TATTCAGCTT	CTCTGGGACATTTCTGFTCAGCAACCCAGTGTGG	6408
6001	GTATGACATCG	TGAAAAATGG	TATTCAGCTT	CTCTGGGACATTTCTGFTCAGCAACCCAGTGTGG	6060
6409	GCAACAAATGAT	CTTTGAGGAA	CATGGTTT	TAGCGGACACACCGCCCAACACGGCCAC	6468
5061	GCAACAAATGAT	CTTTGAGGAA	CATGGTTT	TAGCGGACACACCGCCCAACACGGCCAC	6120
6469	CCCCATAAGGC	ATAGGCCAAGAC	CAATACCCGCGAAT	TATAGGACAAGAGCTCTCTCTCA	6528
6121	CCCCATAAGGC	ATAGGCCAAGAC	CAATACCCGCGAAT	TATAGGACAAGAGCTCTCTCTCA	6180
6529	GACAAACCAT	CTCATGGCCCCCAT	TTCCAGGACACTT	CTGAGTACATCATTTTCATGTCATCC	6588
6181	GACAAACCAT	CTCATGGCCCCCAT	TTCCAGGACACTT	CTGAGTACATCATTTTCATGTCATCC	6240
6589	TGTTGGCATCG	ATGAAGAACCTT	TACGTTCAGGTTCT	CGGAACTTCTACGAGTGGCCAC	6648
6241	TGTTGGCATCG	ATGAAGAACCTT	TACGTTCAGGTTCT	CGGAACTTCTCTACCAAGTGGCCAC	6300
6649	TCCTGACAGG	CCCTCACACAGAGGT	CCACTCAACATCAT	TAGTGGAGGCACTGGAAGACCA	6708
6301	TCCTGACAGG	CCCTCACACAGAGGT	CCACTCAACATCAT	TAGTGGAGGCACTGGAAGACCA	6360
6709	GCAGAGCATAA	GGTTCTGGGAAGAGG	TTTATCCGTGGGCAAT	CTCTGTCAACGAAGCCTT	6768
6361	GCAGAGCATAA	GGTTCTGGGAAGAGG	TTTATCCGTGGGCAAT	CTCTGTCAACGAAGCCTT	6420
6769	GAACCAACCT	ACCGATGACTCGT	CTTTGACCCCTACA	CAGAGTGTCCCATTTATGCCGTGG	6828
6421	GAACCAACCT	ACCGATGACTCGT	CTTTGACCCCTACA	CAGAGTGTCCCATTTATGCCGTGG	6480
6829	AGATGATGGAA	CGAATGTCGNA	TCAGGCTTTAA	ACTCTGTGGCAGTGCTTAGGCTT	6888
6481	AGATGATGGAA	CGAATGTCGNA	TCAGGCTTTAA	ACTCTGTGGCAGTGCTTAGGCTT	6540
6889	TGGAAGTGGT	CATTTCAGATGTGA	TTTATCTAGATGGT	GCATCAATGGTGTGAACATA	6948
6541	TGGAAGTGGT	CATTTCAGATGTGA	TTTATCTAGATGGT	GCATCAATGGTGTGAACATA	6600
6949	CAGATTTGGAG	AGAGTGGGACCGT	CTAGGAGAAATGGCC	ACAGATGATGAGCTGCACATG	7008
6601	CAGATTTGGAG	AGAGTGGGACCGT	CTAGGAGAAATGGCC	ACAGATGATGAGCTGCACATG	6660
7009	TCCTTGGAA	ACGAAAGAGAA	TTCAAGTGTGAC	CCCTCATAGGCAACGTTGTACGATGA	7068
6661	TCCTTGGAA	ACGAAAGAGAA	TTCAAGTGTGAC	CCCTCATAGGCAACGTTGTACGATGA	6720
7069	TGGGACAGCAT	ATACCACTAGGAGAA	CAAGTGGCAGAGGAATAT	CTCGGTGCCATTTTGTCTC	7128
6721	TGGGAGACATA	CCACGTAGGAGAA	CAAGTGGCAGAGGAATAT	CTCGGTGCCATTTTGTCTC	6780
7129	CTGCACATGCT	TTTGGAGCCACG	CGGGGCTGGCGTGTGA	CAATGCGCCGACACTCTGGGGG	7188
6781	CTGCACATGCT	TTTGGAGCCACG	CGGGGCTGGCGTGTGA	CAATGCGCCGACACTCTGGGGG	6840
7189	TGAACCCAGT	CCCCGAGGCA	CTACTGCGCAGT	CCCTCAACACAGTATTTCTCAGAGATACCA	7248
6841	TGAACCCAGT	CCCCGAGGCA	CTACTGCGCAGT	CCCTCAACACAGTATTTCTCAGAGATACCA	6900
7249	TCAGAGAA	CAAAACCTAAT	TTGTAATTTGCC	CAATTTAGTGTCTTATGCTTTAGATGATACA	7308
6901	TCAGAGAA	CAAAACCTAAT	TTGTAATTTGCC	CAATTTAGTGTCTTATGCTTTAGATGATACA	6960

## RESULT 10

```

US-10-236-031B-69
; Sequence 69, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullaus, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B08001/70285 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-236-031B-69

```

Query Match	Score 7423	DB 15	Length 7680
Best Local Similarity	95.2%		
Pred. No. 0	99.8%		
Mismatches 15	Indels 0		
Matches 7432	Conservative		
0	Mismatches 15		
Indels 0			
Gaps 0			

Qy	349	GAAGAGCAAGAGCAGGCTCAGAAATGGTTACGCCCCAGTCCCGGTGCTGTCA	408
Db	1	GAAGAGCAAGAGCAGGCTCAGAAATGGTTACGCCCCAGTCCCGGTGCTGTCA	60
Qy	409	AAGCAAGCCCGTGTATTGACAAATGGAAACACTATCAGATAAAATCAACAGTGGGAGCG	468



2629	Y	TGTAATGTCTATCAGATATCTGAGATGGGAGCAGAGTTTGATCTCTGTCTACTTCA	2688
2281	b	TGTAATGTCTATCAGATATCTGAGATGGGAGCAGAGTTTGATCTCTGTCTACTTCA	2340
2689	Y	AACACAGCGCTGATGCCCCCTCTGACCCCGACTGGACCAAGTTGATGACACCTCAAT	2748
2341	b	AACAAACAGCGCTGATGCCCCCTCTGACCCCGACTGGACCAAGTTGATGACACCTCAAT	2400
2749	Y	TGTTGTTCCGCTGGAGCAGACCCCGAGCTCCCATCAGGGTACAGAAATGTTATTCGCC	2808
2401	b	TGTTGTTCCGCTGGAGCAGACCCCGAGCTCCCATCAGGGTACAGAAATGTTATTCGCC	2460
2809	Y	ATCAGTAGAAGTAGCAGACAGAACTCAACCTTCCTGAACTGCAAACTCCGTCACCT	2868
2461	b	ATCAGTAGAAGTAGCAGACAGAACTCAACCTTCCTGAACTGCAAACTCCGTCACCT	2520
2869	Y	CAGTGACATTGCAACCTGGTGTTCAGTATAACATCACTATCTATGTTGGAAAGAAATCA	2928
2521	b	CAGTGACATTGCAACCTGGTGTTCAGTATAACATCACTATCTATGTTGGAAAGAAATCA	2580
2929	Y	AGAAAGTAGACCTGTTGTCAATTCAACAGAAACCACTGSCACCCACGCTCAGATACAGT	2988
2581	b	AGAAAGTAGACCTGTTGTCAATTCAACAGAAACCACTGSCACCCACGCTCAGATACAGT	2640
2989	Y	GCCCTCTCCACGGACCTGCAGTTTGTGAAAGTGACAGACGTTGAAGGTCAACATCATGTG	3048
2641	b	GCCCTCTCCACGGACCTGCAGTTTGTGAAAGTGACAGACGTTGAAGGTCAACATCATGTG	2700
3049	Y	GACACCGCTGAGAGTGCAGTACCGGCTACCGTGTGGATGTATCCCCGTCAACTGCC	3108
2701	b	GACACCGCTGAGAGTGCAGTACCGGCTACCGTGTGGATGTATCCCCGTCAACTGCC	2760
3109	Y	TGCGCAGACGGGCAGAGGCTGCCCATCAGCAGGAACACCTTTGCAAGAGTACCCGGGCT	3168
2761	b	TGCGCAGACGGGCAGAGGCTGCCCATCAGCAGGAACACCTTTGCAAGAGTACCCGGGCT	2820
3169	Y	GTCCCTCGGGTCACTTATCTTAAAGTCTTTGAGTGAGCCATGGGAGGAGAGCAA	3228
2821	b	GTCCCTCGGGTCACTTATCTTAAAGTCTTTGAGTGAGCCATGGGAGGAGAGCAA	2880
3229	Y	GCCTCTGACTGCTCAACAGACACCAACTGGATGCTCCCACTAACTCCAGTTTCCTCAA	3288
2881	b	GCCTCTGACTGCTCAACAGACACCAACTGGATGCTCCCACTAACTCCAGTTTCCTCAA	2940
3289	Y	TGAAATGATTTCTACTGTCTGTGTGAGATGGATCCACTCGGGCCAGATACAGGATA	3348
2941	b	TGAAATGATTTCTACTGTCTGTGTGAGATGGATCCACTCGGGCCAGATACAGGATA	3000
3349	Y	CCGACTGACCGTGGGCTTTACCCGAGAGAGCAGCCAGCAGTACAAATGTGGTCCCTC	3408
3001	b	CCGACTGACCGTGGGCTTTACCCGAGAGAGCAGCCAGCAGTACAAATGTGGTCCCTC	3060
3409	Y	TGTCTCCAAGTACCCACTGAGGAATCTGACGCTGTCATCTGAGTACACCGTATCCCTCGT	3468
3061	b	TGTCTCCAAGTACCCCTCGAGGAATCTGACGCTGTCATCTGAGTACACCGTATCCCTCGT	3120
3469	Y	GGCCATAAAGGGCAACAGAGAGCCCAAGCCACTGGAGTCTTTACCACTGAGGCC	3528
3121	b	GGCCATAAAGGGCAACAGAGAGCCCAAGCCACTGGAGTCTTTACCACTGAGGCC	3180
3529	Y	TGGGAGCTCTATTTCACCTTACAAACCGAGGTGACTGAGACCAACATTTGATCACATG	3588
3181	b	TGGGAGCTCTATTTCACCTTACAAACCGAGGTGACTGAGACCAACATTTGATCACATG	3240
3589	Y	GACGCTGCTCCAAGAAATTTGTTTTAAGCTGTGGTGTACGACCAAGCCAGGGAGAGGC	3648
3241	b	GACGCTGCTCCAAGAAATTTGTTTTAAGCTGTGGTGTACGACCAAGCCAGGGAGAGGC	3300
3649	Y	ACCACAGAGTGCATTCAGACTCAGGAAGCATGTTGTCTCGGCTTGACTCCAGAGT	3708
3301	b	ACCACAGAGTGCATTCAGACTCAGGAAGCATGTTGTCTCGGCTTGACTCCAGAGT	3360
3709	Y	AGAAATAGCTACACCATCCAAGTCTCTGAGAGATGGACAGAAAGAGATCGCCAAATGT	3768

Db	3361	AGAAATACGCTTACACCATCAAGTCTCTGAGAGATGGACAGGAAAGACATGGCGCAATTGT	3422
Qy	3769	AAACAAAGTGGTGACACCAATTGTTCTCCACCAACAACTTTGCATCTCGAGGCAAAACCTCTGA	3828
Db	3421	AAACAAAGTGGTGACACCAATTGTTCTCCACCAACAACTTTGCATCTCGAGGCAAAACCTCTGA	3480
Qy	3829	CATCTGAGTGCTCACAGTCTCTCGGAGAGGAGCACCCACGAGACATTACTGGTTATAG	3888
Db	3481	CATCTGAGTGCTCACAGTCTCTCGGAGAGGAGCACCCACGAGACATTACTGGTTATAG	3540
Qy	3889	AAATTACCAAAACCCCTACAAAACGCGCAGCAGGAGAAATTTCTTTGGAAAGATGTGTCCTCATGC	3948
Db	3541	AAATTACCAAAACCCCTACAAAACGCGCAGCAGGAGAAATTTCTTTGGAAAGATGTGTCCTCATGC	3600
Qy	3949	TGATCAGAGCTCCTGCACATTTTGATTAACCTGAGTCCGGGCTTGGAGTACAAATCTCAGTGT	4008
Db	3601	TGATCAGAGCTCCTGCACATTTTGATTAACCTGAGTCCGGGCTTGGAGTACAAATCTCAGTGT	3660
Qy	4009	TTACACTGTCAAGGATGACAAAGAAAGTGTCCCTACTCTGTGATACCATCATCCACGAGCTGT	4068
Db	3661	TTACACTGTCAAGGATGACAAAGAAAGTGTCCCTACTCTGTGATACCATCATCCACGAGCTGT	3720
Qy	4069	TCCTCTCCCACTGACCTGGGATTCACCAACATTTGGTTCAGACACCATGGTGTCACTTG	4128
Db	3721	TCCTCTCCCACTGACCTGGGATTCACCAACATTTGGTTCAGACACCATGGTGTCACTTG	3780
Qy	4129	GGCTCCACCCCACTCAATTGATTTAAACCAACTTCTCGTGCCTTACTCACTGTGAACAAA	4188
Db	3781	GGCTCCACCCCACTCAATTGATTTAAACCAACTTCTCGTGCCTTACTCACTGTGAACAAA	3840
Qy	4189	TGAGGAAGATGTTGCCAGAGTGTCAATTTCTCCTTCAGACAATTCGAGTGTCTTTAACAAA	4248
Db	3841	TGAGGAAGATGTTGCCAGAGTGTCAATTTCTCCTTCAGACAATTCGAGTGTCTTTAACAAA	3900
Qy	4249	TCCTCTGCCTGGTACAGAAATATGTAGTGAAGTGTCTCCAGTGTCTACGAAACAAATGAGAG	4308
Db	3901	TCCTCTGCCTGGTACAGAAATATGTAGTGAAGTGTCTCCAGTGTCTACGAAACAAATGAGAG	3960
Qy	4309	CACACCTCTTAGAGGAAGACAGAAAAACAGTCTTGATTTCCCCAACTGGCAATTGACTTTTC	4368
Db	3961	CACACCTCTTAGAGGAAGACAGAAAAACAGTCTTGATTTCCCCAACTGGCAATTGACTTTTC	4020
Qy	4369	TGATATTACTGCCAACTCTTTTACTGTGCACTGGATTGCTCTCGAGCCACCATCACTGG	4428
Db	4021	TGATATTACTGCCAACTCTTTTACTGTGCACTGGATTGCTCTCGAGCCACCATCACTGG	4080
Qy	4429	CTACAGGATCGCCCATCATCCGAGCACTTCAGTGGGAGACCTCGAGAAGATCGSGTGCC	4488
Db	4081	CTACAGGATCGCCCATCATCCGAGCACTTCAGTGGGAGACCTCGAGAAGATCGSGTGCC	4140
Qy	4489	CCACTCTCGGAATTCATCACTCCAGCACTCATCCAGCACTCATCCAGGCAAGATGTGGTCAG	4548
Db	4141	CCACTCTCGGAATTCATCACTCCAGCACTCATCCAGCACTCATCCAGGCAAGATGTGGTCAG	4200
Qy	4549	CATCGTTGCTCTTAATGGCAGAGAGAAAGTCCCTTATGTATTGGCCAAACAAATCAACAGT	4608
Db	4201	CATCGTTGCTCTTAATGGCAGAGAGAAAGTCCCTTATGTATTGGCCAAACAAATCAACAGT	4260
Qy	4609	TTCTGATGTTCCGAGGACCTGGAAAGTTGTGTGCGACCCCAACAGCCTTACTGTATCAG	4668
Db	4261	TTCTGATGTTCCGAGGACCTGGAAAGTTGTGTGCGACCCCAACAGCCTTACTGTATCAG	4320
Qy	4669	CTGGGATGCTCTCCTGTGCAGTGGAGATATACAGGATCACTTTACGGAGAAAACAGAGG	4728
Db	4321	CTGGGATGCTCTCCTGTGTGCAGTGGAGATATACAGGATCACTTTACGGAGAAAACAGAGG	4380
Qy	4729	AAATAGCCCTGTCCAGGAGTTCACTGTCCCTGGGAGCAAGTCTACAGCTACCATCAGCGG	4788
Db	4381	AAATAGCCCTGTCCAGGAGTTCACTGTCCCTGGGAGCAAGTCTACAGCTACCATCAGCGG	4440
Qy	4789	CCTTTAAACCTGGAGTTGATTAACCATCACTGTGTATGCTGTCTACTGGCCGTGGAGACAG	4848

b 4441 CCTTAAACCTGGAGTGTATTATACCATCACTGTGTATGCTGTCACTGGCGGTGGAGACAG 4500  
 y 4849 CCCCAGACAGCAAGCCAAATTCATTAATTAACGAAACAGAAATTAACAACCATCCCA 4908  
 b 4501 CCCCAGACAGCAAGCCAAATTCATTAATTAACGAAACAGAAATTAACAACCATCCCA 4560  
 y 4909 GATCAAGTGCAGGATGTTTCAGGACCAACAGCATTAAGTGTCAAGTGGCTGCTTCAAGTTC 4968  
 b 4561 GATCAAGTGCAGGATGTTTCAGGACCAACAGCATTAAGTGTCAAGTGGCTGCTTCAAGTTC 4620  
 y 4969 CCTGTGTACTGTGTACAGAGTAAACCACTCCCAAAATTAAGACAGACCAACAAAAAC 5028  
 b 4621 CCTGTGTACTGTGTACAGAGTAAACCACTCCCAAAATTAAGACAGACCAACAAAAAC 4680  
 y 5029 TAAACTGCAGGTCCAGATCAACAGAAATGACTATTAAGGCTTGCAGGCCACAGATGGA 5088  
 b 4681 TAAACTGCAGGTCCAGATCAACAGAAATGACTATTAAGGCTTGCAGGCCACAGATGGA 4740  
 y 5089 GTATGTGGTGTAGTGTCTATGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGGTTCAGAC 5148  
 b 4741 GTATGTGGTGTAGTGTCTATGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGGTTCAGAC 4800  
 y 5149 TGCAGTACCAACATTAAGTGCCTTAAGGACTGGCATTCACTGATGTGGATGTCGATTC 5208  
 b 4801 TGCAGTACCAACATTAAGTGCCTTAAGGACTGGCATTCACTGATGTGGATGTCGATTC 4860  
 y 5209 CATCAAAATGTCTTGGAAAGCCCAAGGGGCAAGTTTCCAGGTACAGGGTGAACCTACTC 5268  
 b 4861 CATCAAAATGTCTTGGAAAGCCCAAGGGGCAAGTTTCCAGGTACAGGGTGAACCTACTC 4920  
 y 5269 GAGCCCTGAGATGGAATCCAGTATGCTATCCCTGACCTGATGTGGATGTCGATTC 5328  
 b 4921 GAGCCCTGAGATGGAATCCAGTATGCTATCCCTGACCTGATGTGGATGTCGATTC 4980  
 y 5329 AGAGCTGCAAGGCTCAGACCGGGTCTCAGTACACAGTCAAGTGTGGTTCCTTGCACGA 5388  
 b 4981 AGAGCTGCAAGGCTCAGACCGGGTCTCAGTACACAGTCAAGTGTGGTTCCTTGCACGA 5040  
 y 5389 TGATATGGAGAGCCAGCCCTGATTTGGAAACCGAGTCCAGAGTATTCCTGACACCACTGA 5448  
 b 5041 TGATATGGAGAGCCAGCCCTGATTTGGAAACCGAGTCCAGAGTATTCCTGACACCACTGA 5100  
 y 5449 CCTGAAGTTCACCTCAGTCAACCCAGAGCTGAGCGCCAGTGGACACCAACCAATGT 5508  
 b 5101 CCTGAAGTTCACCTCAGTCAACCCAGAGCTGAGCGCCAGTGGACACCAACCAATGT 5160  
 y 5509 TCAGCTCAGTGGATATCGAGTGGGGTGAACCCCAAGAGAGAACCGGACCAATGAAGA 5568  
 b 5161 TCAGCTCAGTGGATATCGAGTGGGGTGAACCCCAAGAGAGAACCGGACCAATGAAGA 5220  
 y 5569 AATCAACCTTGTCTCAGAGCTCATCCGTGGTGTATCAGAGCTTATGGTGGCCACCA 5628  
 b 5221 AATCAACCTTGTCTCAGAGCTCATCCGTGGTGTATCAGAGCTTATGGTGGCCACCA 5280  
 y 5629 ATATGAAGTGAAGTGTCTATGCTCTTAAGGACATTTGAACAGCAGACCAAGCTCAGGGAGT 5688  
 b 5281 ATATGAAGTGAAGTGTCTATGCTCTTAAGGACATTTGAACAGCAGACCAAGCTCAGGGAGT 5340  
 y 5689 TGTCAACCTTGTCTCAGAGCTCATCCGTGGTGTATCAGAGCTTATGGTGGCCACCA 5748  
 b 5341 TGTCAACCTTGTCTCAGAGCTCATCCGTGGTGTATCAGAGCTTATGGTGGCCACCA 5400  
 y 5749 GACCAACCATCAACATTAAGTGTGAGAACCAAGACTGAGACGATCACTGGCTTCAAGTTGA 5808  
 b 5401 GACCAACCATCAACATTAAGTGTGAGAACCAAGACTGAGACGATCACTGGCTTCAAGTTGA 5460  
 y 5809 TGGGTTCCAGCCAAATGGCCAGACTCCATCCAGAGAACCAATCAAGCCAGATGTCAGAG 5868  
 b 5461 TGGCGTTCCAGCCAAATGGCCAGACTCCATCCAGAGAACCAATCAAGCCAGATGTCAGAG 5520  
 y 5869 CTACACCATCACTGGCTTACAAACAGGCACTGACTACAGATCTACTGTACACCTTGAA 5928  
 b 5521 CTACACCATCACTGGCTTACAAACAGGCACTGACTACAGATCTACTGTACACCTTGAA 5580

Qy 5929 TGACAAATGCTCGAGTCCCTGTGTGTATCATCGACGCTCCCACTGCTGATGATGACCAATC 5988  
 Db 5581 TGACAAATGCTCGAGTCCCTGTGTGTGTATCATCGACGCTCCCACTGCTGATGATGACCAATC 5640  
 Qy 5989 CAACTGCGTTCCTGGCCACACACCCAAATTCCTTGTGTGTATCATGGCAGCCGCCAG 6048  
 Db 5641 CAACTGCGTTCCTGGCCACACACCCAAATTCCTTGTGTGTATCATGGCAGCCGCCAG 5700  
 Qy 6049 TGCACGATTAACCGGTACATCATCAAGTATCAGAAAGCTTGGTCTCTCTCCAGAGAAAT 6108  
 Db 5701 TGCACGATTAACCGGTACATCATCAAGTATCAGAAAGCTTGGTCTCTCTCCAGAGAAAT 5760  
 Qy 6109 GGTCCCTCGCCCGCCCTGGTGTCAAGAGGCTACTATTAATCTGGCTGGAAACCGGGAAAC 6168  
 Db 5761 GGTCCCTCGCCCGCCCTGGTGTCAAGAGGCTACTATTAATCTGGCTGGAAACCGGGAAAC 5820  
 Qy 6169 CGAATATCAATTTATGTCTATTCCTCGCTGAAGAAATATCAGAAAGCGAGCCCTGATTCG 6228  
 Db 5821 CGAATATCAATTTATGTCTATTCCTCGCTGAAGAAATATCAGAAAGCGAGCCCTGATTCG 5880  
 Qy 6229 AAGGAAAAAGACAGACGAGTTCCTCCAACTGTGTAACTTCCACACCCCAATCTTCATCG 6288  
 Db 5881 AAGGAAAAAGACAGACGAGTTCCTCCAACTGTGTAACTTCCACACCCCAATCTTCATCG 5940  
 Qy 6289 ACAGAGATCTTGGATGTTCTTCCACAGTTCAAAGAGCCCTTTCGTTCACCCACCCCTGG 6348  
 Db 5941 ACCAGAGATCTTGGATGTTCTTCCACAGTTCAAAGAGCCCTTTCGTTCACCCACCCCTGG 6000  
 Qy 6349 GTATGACACTGGAATATGTATTCAGCTTCTGTGCACTTCTGTGTGAGCAACCCAGTGTTCG 6408  
 Db 6001 GTATGACACTGGAATATGTATTCAGCTTCTGTGCACTTCTGTGTGAGCAACCCAGTGTTCG 6060  
 Qy 6409 GCAACAAATGATCTTTGAGGAAACATGTTTAGGCGGACACACCCGCCACACAGCCGAC 6468  
 Db 6061 GCAACAAATGATCTTTGAGGAAACATGTTTAGGCGGACACACCCGCCACACAGCCGAC 6120  
 Qy 6469 CCCCATAAGGCATAGGCCAAGACCATACCCGCCGAATGTAGGCAAGAGCTCTCTCTCA 6528  
 Db 6121 CCCCATAAGGCATAGGCCAAGACCATACCCGCCGAATGTAGGCAAGAGCTCTCTCTCA 6180  
 Qy 6529 GACAACTCATTTGGGCCCCCATTCAGAGCACTTCTGAGTACATCATTTTCATGTCATCC 6588  
 Db 6181 GACAACTCATTTGGGCCCCCATTCAGAGCACTTCTGAGTACATCATTTTCATGTCATCC 6240  
 Qy 6589 TGTGGCACTGATGAAGAACCTTACAGTTCAGGGTTCCTGGAACCTTCTACAGTCCAC 6648  
 Db 6241 TGTGGCACTGATGAAGAACCTTACAGTTCAGGGTTCCTGGAACCTTCTACAGTCCAC 6300  
 Qy 6649 TCTGACAGGCTCAACAGAGTGCACCTACAAATCATAGTGGAGGCACTGAAAGACCA 6708  
 Db 6301 TCTGACAGGCTCAACAGAGTGCACCTACAAATCATAGTGGAGGCACTGAAAGACCA 6360  
 Qy 6709 GCAGAGCATTAAGTTCGGGAAGAGGTTGTACCGTGGGCACTCTGTCAACGAAAGCTT 6768  
 Db 6361 GCAGAGCATTAAGTTCGGGAAGAGGTTGTACCGTGGGCACTCTGTCAACGAAAGCTT 6420  
 Qy 6769 GAACCAACCTACGATGACTCGTGTTCGACCCCTACAGTGTCCCATTAATGCCCCGTTGG 6828  
 Db 6421 GAACCAACCTACGATGACTCGTGTTCGACCCCTACAGTGTCCCATTAATGCCCCGTTGG 6480  
 Qy 6829 AGATGAGTGGAGAGATGCTGAATCAGGCTTTAACTGTCGAGTGTCTTAGGCTT 6888  
 Db 6481 AGATGAGTGGAGAGATGCTGAATCAGGCTTTAACTGTCGAGTGTCTTAGGCTT 6540  
 Qy 6889 TGAAGTGGTTCATTTAGATGTGATTCATCTAGATGGTGCATGACAAATGGTGTGAACCTA 6948  
 Db 6541 TGAAGTGGTTCATTTAGATGTGATTCATCTAGATGGTGCATGACAAATGGTGTGAACCTA 6600  
 Qy 6949 CAAATTTGGAGAGAGTGGGACCGTTCAGGGAGAAATGGCCAGATGATGAGTGCACATG 7008  
 Db 6601 CAAATTTGGAGAGAGTGGGACCGTTCAGGGAGAAATGGCCAGATGATGAGTGCACATG 6660

```

7009 TCTTGGGAACGGAAGAGAAATCAAGTGTGACCTCATGAGCGAACGTTGATGATGA 7068
6661 TCTTGGGAACGGAAGAGAAATCAAGTGTGACCTCATGAGCGAACGTTGATGATGA 6720
7069 TGGGAAGACATACCACTAGGAGAACTGGGAGAGAAATATCTGGTGGCAATTTGCTC 7128
6721 TGGGAAGACATACCACTAGGAGAACTGGGAGAGAAATATCTGGTGGCAATTTGCTC 6780
7129 CTGCACATGCTTTGGAGGCGAGCGGGCTGGCGCTGTGAACAATGCGCGAGACCTGGGG 7188
6781 CTGCACATGCTTTGGAGGCGAGCGGGCTGGCGCTGTGAACAATGCGCGAGACCTGGGG 6840
7189 TGAACCCAGTCCCGAGGCACTACTGCGCAGTCTCAACAGGATTTCTCAGAGATACCA 7248
6841 TGAACCCAGTCCCGAGGCACTACTGCGCAGTCTCAACAGGATTTCTCAGAGATACCA 6900
7249 TCAGAGAACAAACATAATGTTAATGCCCAATGAGTGTCTTATGCGCTTTAGATGACA 7308
6901 TCAGAGAACAAACATAATGTTAATGCCCAATGAGTGTCTTATGCGCTTTAGATGACA 6960
7309 GGCTGACAGAGAGATTCGCGAGGATTAATCATCTTTCCCAATCCAGAGGAACAGCATGT 7368
6961 GGCTGACAGAGAGATTCGCGAGGATTAATCATCTTTCCCAATCCAGAGGAACAGCATGT 7020
7369 CTCTCTGCCAAGATCCATCTAAACTGAGGATGATGTAGCAGACCCAGCTTTAGAGTTCTTC 7428
7021 CTCTCTGCCAAGATCCATCTAAACTGAGGATGATGTAGCAGACCCAGCTTTAGAGTTCTTC 7080
7429 TTTCTTTCTTAAGCCCTTGTCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCAGAGCT 7488
7081 TTTCTTTCTTAAGCCCTTGTCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCAGAGCT 7140
7489 TCTCCAAAGCATCACCTCGGAGTTTCTCTGAGGGTTTCTCATAAATGAGGCGCTGCACATT 7548
7141 TCTCCAAAGCATCACCTCGGAGTTTCTCTGAGGGTTTCTCATAAATGAGGCGCTGCACATT 7200
7549 GCCTGTCTTGTGTCGAGATTTCAATACCGCTCAGTATTTTAAATGAGTATCTTAAGA 7608
7201 GCCTGTCTTGTGTCGAGATTTCAATACCGCTCAGTATTTTAAATGAGTATCTTAAGA 7260
7609 TTTGGTTTGGGATCAATAGGAAGCATATGAGCAACCAAGATGCAAAATGTTTGAAT 7668
7261 TTTGGTTTGGGATCAATAGGAAGCATATGAGCAACCAAGATGCAAAATGTTTGAAT 7320
7669 GATATGACCAAAATTTAAGTAGGAAGTCAACCAACACTTCTGCTTTCACTTAAGTGT 7728
7321 GATATGACCAAAATTTAAGTAGGAAGTCAACCAACACTTCTGCTTTCACTTAAGTGT 7380
7729 CTGGCCGCAATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTAAATATCCA 7788
7381 CTGGCCGCAATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTAAATATCCA 7440
7789 CAGTACT 7795
7441 CAGTACT 7447

```

RESULT 11

```

US-10-374-979-75
; Sequence 75, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13

```

```

; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 75
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-374-979-75

Query Match          95.2%; Score 7423; DB 15; Length 7680;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

349  GAAGAGCAAGAGGAGGCTCAGCAANTGGTTACGCCCCAGTCCCGGTGGGTGCTGCAGTCA 408
Db      1  GAAGAGCAAGAGGAGGCTCAGCAANTGGTTACGCCCCAGTCCCGGTGGGTGCTGCAGTCA 60

409  AAGCAAGCCCGGTTGTTATGACAAATGGAAAAACACTATCAGATAAAATCAACAGTGGAGCG 468
Db      61  AAGCAAGCCCGGTTGTTATGACAAATGGAAAAACACTATCAGATAAAATCAACAGTGGAGCG 120

469  GACCTACCTAGCAATGCGTTGGTTTGTACTTGTATGGAGAAAGCCGAGGTTTAACTG 528
Db      121  GACCTACCTAGGTAATGTGTTGGTTGTACTTGTATGGAGAAAGCCGAGGTTTAACTG 180

529  CCAAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTGACAAAGTACACTGGGAACACTTTACCG 588
Db      181  CCAAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTGACAAAGTACACTGGGAACACTTTACCG 240

589  AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTTGCATCGG 648
Db      241  AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTTGCATCGG 300

649  GCCTGGGAGGAGGAGAAATGAAGCTGACCAATCGAAACCGCTGCCATGAAGGGGGTCACTC 708
Db      301  GCCTGGGAGGAGGAGAAATGAAGCTGACCAATCGAAACCGCTGCCATGAAGGGGGTCACTC 360

709  CTACAAGATTTGGTGCACCTGGAGAGACCAATGAGACTCGGTGTTTACATGTTTAGAGTG 768
Db      361  CTACAAGATTTGGTGCACCTGGAGAGACCAATGAGACTCGGTGTTTACATGTTTAGAGTG 420

769  TGTGTGCTCTGTTAATGGAAGAGGAAATGGACCTGCAAGCCCATAGCTGAGAGAGTGT 828
Db      421  TGTGTGCTCTGTTAATGGAAGAGGAAATGGACCTGCAAGCCCATAGCTGAGAGAGTGT 480

829  TGATCATGCTGCTGGGACTTCCCTATGTGTGGTGGGAGAAACGTTGGGAGAGCCCTACCAAG 888
Db      481  TGATCATGCTGCTGGGACTTCCCTATGTGTGGTGGGAGAAACGTTGGGAGAGCCCTACCAAG 540

889  CTGGATGATGGTAGATTGTACTTGGCTGGGAGAGGAGCGGACGCGCATCACTTGCACATTC 948
Db      541  CTGGATGATGGTAGATTGTACTTGGCTGGGAGAGGAGCGGACGCGCATCACTTGCACATTC 600

949  TAGAAATAGATCAACAGGATCAGGACACCAAGGACATCTCTATAGAATTGGAGACACCTGGAG 1008
Db      601  TAGAAATAGATCAACAGGATCAGGACACCAAGGACATCTCTATAGAATTGGAGACACCTGGAG 660

1009  CAAGAAGGATTAATCAGGAAACCTCTCTCAGTGTACTGTCAGAGGAGGAGCGGAGGAGA 1068
Db      661  CAAGAAGGATTAATCAGGAAACCTCTCTCAGTGTACTGTCAGAGGAGGAGCGGAGGAGA 720

1069  GTGGAAGTGTGAGAGGACACCTCTGTGCAGACCACTCATCGAGCGGATCTGGGCCCTTCAAC 1128
Db      721  GTGGAAGTGTGAGAGGACACCTCTGTGCAGACCACTCATCGAGCGGATCTGGGCCCTTCAAC 780

1129  CGATGTTCTGTGAGTGTGTTTAAACCGGAGCGCTCAACCCGAGCGCTCTCTCCCTATGGCA 1188
Db      781  CGATGTTCTGTGAGTGTGTTTAAACCGGAGCGCTCAACCCGAGCGCTCTCTCCCTATGGCA 840

1189  CTGTGTCACAGACAGTGGTGTCTTACTCTGTGGGATGCAAGTGGCTGAAACACACAAGG 1248

```



b 841 CTGTGTACAGACAGTGTGTGTCTACTCTCTGTGGGATGAGTGGTTGAAGACACAAAGG 900  
y 1249 AAATAAGCAAAATGCTTTTGACCTGCTGCTGGCAACCGAGTCTAGCTGCCAAGAGACAGCTGT 1308  
b 901 AAATAAGCAAAATGCTTTTGACCTGCTGCTGGCAACCGAGTCTAGCTGCCAAGAGACAGCTGT 960  
y 1309 RACCAGACTTACGCTGGCAACTCAATGAGAGCCATGTGCTTTACCAATTCACCTACAA 1368  
b 961 AACCCAGACTTACGCTGGCAACTTAAATGAGAGCCATGTGCTTTACCAATTCACCTACAA 1020  
y 1369 TGGCAGGACGTTTACTCTGACCAACAGAGGGGAGAGGAGCAATCTTTGGTGGAG 1428  
b 1021 TGGCAGGACGTTTACTCTGACCAACAGAGGGGAGAGGAGCAATCTTTGGTGGAG 1080  
y 1429 CACAACTTCGAATATGACGAGCAACAGAAATACCTTTCTGACAGACCAACACTGTTTT 1488  
b 1081 CACAACTTCGAATATGACGAGCAACAGAAATACCTTTCTGACAGACCAACACTGTTTT 1140  
y 1489 GGTTACAGACTCGAGGAGGAAATTCGAATGGTGCCTGTGACCACTTCCCTTCCTATAAA 1548  
b 1141 GGTTACAGACTCGAGGAGGAAATTCGAATGGTGCCTGTGACCACTTCCCTTCCTATAAA 1200  
y 1549 CAACCAATTAACACTGATGACCTCTGAGGGAGAGGAGACACATCAAGTGGTGG 1608  
b 1201 CAACCAATTAACACTGATGACCTCTGAGGGAGAGGAGACACATCAAGTGGTGG 1260  
y 1609 GACCAACACAACTATGATGCGGACAGCAAGTTTGGGTTCTGCCCCCTGGCTGCCACGA 1668  
b 1261 GACCAACACAACTATGATGCGGACAGCAAGTTTGGGTTCTGCCCCCTGGCTGCCACGA 1320  
y 1669 GGAAATCTGCACAAACCAATGAAGGGTCTATGACCGATGAGATGAGATCACTTACAAATGT 1728  
b 1321 GGAAATCTGCACAAACCAATGAAGGGTCTATGACCGATGAGATGAGATCACTTACAAATGT 1380  
y 1729 GCATGACATGGGTTCACATGATGAGGTGCAGTGTGTTGGGAATGGTCTGGGGGAATGGAC 1788  
b 1381 GCATGACATGGGTTCACATGATGAGGTGCAGTGTGTTGGGAATGGTCTGGGGGAATGGAC 1440  
y 1789 ATGCATGTCCTACTCCGAGTTGAGATCAGTGCATATGTTGATGACATCACTTACAAATGT 1848  
b 1441 ATGCATGTCCTACTCCGAGTTGAGATCAGTGCATATGTTGATGACATCACTTACAAATGT 1500  
y 1849 GAACGACATTCACAAAGCGTTCATGAAGGGGACATGCTGAACTGTACATGCTTGG 1908  
b 1501 GAACGACATTCACAAAGCGTTCATGAAGGGGACATGCTGAACTGTACATGCTTGG 1560  
y 1909 TCAGGGTCGGGGCAGGTGGAAGTGTGATCCGTCGACCAATGCCAGGATTCAGAGACTGG 1968  
b 1561 TCAGGGTCGGGGCAGGTGGAAGTGTGATCCGTCGACCAATGCCAGGATTCAGAGACTGG 1620  
y 1969 GACGTTTTATCAAAATGGAGATTCATGGGAGAGTATGTGCATGGTGCAGATACCAAGTG 2028  
b 1621 GACGTTTTATCAAAATGGAGATTCATGGGAGAGTATGTGCATGGTGCAGATACCAAGTG 1680  
y 2029 CTACTGCTATGGCCGTGGCAATGGGAGTGGCAATGCCAACCTTTACAGACTATCCAAAG 2088  
b 1681 CTACTGCTATGGCCGTGGCAATGGGAGTGGCAATGCCAACCTTTACAGACTATCCAAAG 1740  
y 2089 CTCAGTGGTCTCTCGAAGTATTTATCTAGACTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 2148  
b 1741 CTCAGTGGTCTCTCGAAGTATTTATCTAGACTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 1800  
y 2149 CATCCAGTGGAAATGACCAAGCCATCTCAATTTCCAAAGTACATTTCTCAGTGGAGACC 2208  
b 1801 CATCCAGTGGAAATGACCAAGCCATCTCAATTTCCAAAGTACATTTCTCAGTGGAGACC 1860  
y 2209 TAAATATCTGAGGCGGTGGAGGAGCTACATACAGGCACTTAACTCTCTACAC 2268  
b 1861 TAAATATCTGAGGCGGTGGAGGAGCTACATACAGGCACTTAACTCTCTACAC 1920  
y 2269 CATCAAGGCGCTGAAGCCTGGTGTATACGAGGGCCAGCTCATCAGCATCCAGCAGTA 2328  
b 1921 CATCAAGGCGCTGAAGCCTGGTGTATACGAGGGCCAGCTCATCAGCATCCAGCAGTA 1980

Qy 2329 CGGCCACCAAGAGTCACTCGCTTTGACTTCAACCACACAGACACACCTGTGTAC 2388  
Db 1981 CGGCCACCAAGAGTCACTCGCTTTGACTTCAACCACACAGACACACCTGTGTAC 2040  
Qy 2389 CAGCAACACCGTGACAGGAGAGCACTCCCTTTTCTCTCTTGTGGCCACTTCTGAATC 2448  
Db 2041 CAGCAACACCGTGACAGGAGAGCACTCCCTTTTCTCTCTTGTGGCCACTTCTGAATC 2100  
Qy 2449 TGTGACCGAAATCAGACAGCAGTAGCTTTGTGCTCTCTGCTCTCAGCTTCCGACACCGT 2508  
Db 2101 TGTGACCGAAATCAGACAGCAGTAGCTTTGTGCTCTCTGCTCTCAGCTTCCGACACCGT 2160  
Qy 2509 GTCGGATTCGGGTTGGAATATGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGA 2568  
Db 2161 GTCGGATTCGGGTTGGAATATGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGA 2220  
Qy 2569 TCTTCNAGACAGACCACTCTGTGAACATCCCTGACCTGCTCTCTGGCGGAAATACAT 2628  
Db 2221 TCTTCNAGACAGACCACTCTGTGAACATCCCTGACCTGCTCTCTGGCGGAAATACAT 2280  
Qy 2629 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTTTGTCTCTACTTTACA 2688  
Db 2281 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTTTGTCTCTACTTTACA 2340  
Qy 2689 AACAAAGCCCTGTATGCCCTCTCTGACCGGCTGTGGACCAAGTTTGTATGACACCTCAAT 2748  
Db 2341 AACAAAGCCCTGTATGCCCTCTCTGACCGGCTGTGGACCAAGTTTGTATGACACCTCAAT 2400  
Qy 2749 TGTGTTGCTGTGAGACAGACCCCGAGCTCCCATCACAGGGGTACAGAAATAGTCTATTGCGC 2808  
Db 2401 TGTGTTGCTGTGAGACAGACCCCGAGCTCCCATCACAGGGGTACAGAAATAGTCTATTGCGC 2460  
Qy 2809 ATCAGTAGAAGGTAGACGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCT 2868  
Db 2461 ATCAGTAGAAGGTAGACGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCT 2520  
Qy 2869 CAGTGACTTCCAACTGCTTCAAGTATTAACATCACTATCTATGCTGTGGAAGAAATCA 2928  
Db 2521 CAGTGACTTCCAACTGCTTCAAGTATTAACATCACTATCTATGCTGTGGAAGAAATCA 2580  
Qy 2929 AGAAAGTACACCTGTTGTCTATTCAAAGAAACCACTGGCACCCGCTCAGATACAGT 2988  
Db 2581 AGAAAGTACACCTGTTGTCTATTCAAAGAAACCACTGGCACCCGCTCAGATACAGT 2640  
Qy 2989 GCCCTCTCCAGGAGCTGACGTTTGTGGAAGTACAGACGCTGAAGTCAACCTCATATGT 3048  
Db 2641 GCCCTCTCCAGGAGCTGACGTTTGTGGAAGTACAGACGCTGAAGTCAACCTCATATGT 2700  
Qy 3049 GACACCGCTGTGAGAGTGCAGTACCGGCTACCGGTGAGTGTGATCCCGTCAACCTGCC 3108  
Db 2701 GACACCGCTGTGAGAGTGCAGTACCGGCTACCGGTGAGTGTGATCCCGTCAACCTGCC 2760  
Qy 3109 TGGGAGCAGCGGGCAGAGGCTGCCATCAGCAGAAACCTTTGCAAGAGTCAACCGGCT 3168  
Db 2761 TGGGAGCAGCGGGCAGAGGCTGCCATCAGCAGAAACCTTTGCAAGAGTCAACCGGCT 2820  
Qy 3169 GTCCCTCTGGGGTCACTTATTACTTCAAAGTCTTTGCAAGTCAAGGAGGAGAGCAA 3228  
Db 2821 GTCCCTCTGGGGTCACTTATTACTTCAAAGTCTTTGCAAGTCAAGGAGGAGAGCAA 2880  
Qy 3229 GCCTCTGACTGCTCAACAGACAAACCTGGATGTCCCACTAACTCCAGTTTGTCAA 3288  
Db 2881 GCCTCTGACTGCTCAACAGACAAACCTGGATGTCCCACTAACTCCAGTTTGTCAA 2940  
Qy 3289 TGAACCTGATTCTACTGCTCTGTGATGGAATCCACCTCGGGCCCGAGATAACAGGATA 3348  
Db 2941 TGAACCTGATTCTACTGCTCTGTGATGGAATCCACCTCGGGCCCGAGATAACAGGATA 3000  
Qy 3349 CCGACTGACCGTGGGCTTACCCGAGAGAGACAGCCAGGACGATCAATGTGGGTCCCTC 3408  
Db 3001 CCGACTGACCGTGGGCTTACCCGAGAGAGGACAGCCAGGACGATCAATGTGGGTCCCTC 3060





5221 ATCAACCTTGCTCTGACAGCTCATCCGTGTTGTATCAGAGCTTATGTGGCCACAA 5280  
5629 ATATGAAGTGAAGTCTATGTCTTTAAGGACACCTTTGACAGGACAGAGCTCAGGAGT 5688  
5281 ATATGAAGTGAAGTCTATGTCTTTAAGGACACCTTTGACAGGACAGAGCTCAGGAGT 5340  
5689 TGCTCAACCACTCTGGAGATGTGAGCCCAACCAAGAGGCTGTGTGACAGATGCTACTGA 5748  
5341 TGTCAACCACTCTGGAGATGTGAGCCCAACCAAGAGGCTGTGTGACAGATGCTACTGA 5400  
5749 GACCAACCACTCACTATGCTGGAGAACCAAGACTGAGAGATCACTGCGCTCCAGTTGA 5808  
5401 GACCAACCACTCACTATGCTGGAGAACCAAGACTGAGAGATCACTGCGCTCCAGTTGA 5460  
5809 TCCCGTTCCAGCAATGGCCAGACTCCCAATCCAGAGAACCAAGCTGAGAGATGTCAGAG 5868  
5461 TCCCGTTCCAGCAATGGCCAGACTCCCAATCCAGAGAACCAAGCTGAGAGATGTCAGAG 5520  
5869 CTACACCACTCACTGCTTTACACCAAGGCACTGACTACAGAGATCTACCTGTACACCTGAA 5928  
5521 CTACACCACTCACTGCTTTACACCAAGGCACTGACTACAGAGATCTACCTGTACACCTGAA 5580  
5929 TGACCAATGCTCGAGCTCCCTGTGTGATCATGACGCTCCACTGCGCTGATGACCACTC 5988  
5581 TGACCAATGCTCGAGCTCCCTGTGTGATCATGACGCTCCACTGCGCTGATGACCACTC 5640  
5989 CAACCTGCGCTTCTGGCCACCAACCCAAATCTGCTGTGATCATGCGAGCGCCACG 6048  
5641 CAACCTGCGCTTCTGGCCACCAACCCAAATCTGCTGTGATCATGCGAGCGCCACG 5700  
6049 TGCCAGGATTCAGGCTCATCATCAAGTATCAGAGCTGCGCTGCTCCACAGAGAGT 6108  
5701 TGCCAGGATTCAGGCTCATCATCAAGTATCAGAGCTGCGCTGCTCCACAGAGAGT 5760  
6109 GGTCTCTGCGCCCGCTCGTGTCTACAGAGCTACTATTACTGCGCTGGAACCCGGAAC 6168  
5761 GGTCTCTGCGCCCGCTCGTGTCTACAGAGCTACTATTACTGCGCTGGAACCCGGAAC 5820  
6169 CGAATATCAATTTATGTCTTCCAGCTTCCAGAGCTTCCAGAGAGCTGCTCCAGAGAGT 6228  
5821 CGAATATCAATTTATGTCTTCCAGCTTCCAGAGCTTCCAGAGAGCTGCTCCAGAGAGT 5880  
6229 AAGGAAAAAGACAGACGAGCTTCCCAATCGTGTAAACCTTCCACACCCCAATCTTCATGG 6288  
5881 AAGGAAAAAGACAGACGAGCTTCCCAATCGTGTAAACCTTCCACACCCCAATCTTCATGG 5940  
6289 ACCAGAGATCTTGGATGTTCTTCCAGCTTCCAGAGCTTCCAGAGAGCTTCCAGAGAGT 6348  
5941 ACCAGAGATCTTGGATGTTCTTCCAGCTTCCAGAGCTTCCAGAGAGCTTCCAGAGAGT 6000  
6349 GTATGACACTGGAAATGGTATTGAGCTTCTGCGCACTTCTGCTGAGCAACCCAGTGTGG 6408  
6001 GTATGACACTGGAAATGGTATTGAGCTTCTGCGCACTTCTGCTGAGCAACCCAGTGTGG 6060  
6409 GCAACAAATGATCTTTGAGGAACATGGTTTATAGGAGGACCAACCGCCCAACAGCCAC 6468  
6061 GCAACAAATGATCTTTGAGGAACATGGTTTATAGGAGGACCAACCGCCCAACAGCCAC 6120  
6469 CCCATAAGGCATAGGCGCAAGACATACCGCGCGAATGTAGGACAGAGCTCTCTCTCA 6528  
6121 CCCATAAGGCATAGGCGCAAGACATACCGCGCGAATGTAGGACAGAGCTCTCTCTCA 6180  
6529 GACAAACATCTATGGGCCCCATTCAGGACACTTCTGAGTACATCATTTTATGTCATCC 6588  
6181 GACAAACATCTATGGGCCCCATTCAGGACACTTCTGAGTACATCATTTTATGTCATCC 6240  
6589 TGTGTCACCTGATGAAGAACCTTACAGTTCAGGCTTCTGAGAACTTCTACAGTGGCCAC 6648  
6241 TGTGTCACCTGATGAAGAACCTTACAGTTCAGGCTTCTGAGAACTTCTACAGTGGCCAC 6300  
6649 TCTGACAGGCTTCCAGAGGTCGCCACCTAACATCATATGATGGAGGCACTGAAGACCA 6708  
6301 TCTGACAGGCTTCCAGAGGTCGCCACCTAACATCATATGATGGAGGCACTGAAGACCA 6360

QY 6709 GCAGAGCATTAAGTTTCGGGAAGAGTTGTTTACCGTGGGCAACTCTGTCAACGAAGGCTT 6768  
DB 6361 GCAGAGCATTAAGTTTCGGGAAGAGTTGTTTACCGTGGGCAACTCTGTCAACGAAGGCTT 6420  
QY 6769 GAACCAACCTTACGGATGACTCGTGTCTTTGACCCCTTACACAGTGTCCCATTTATGCGTTGG 6828  
DB 6421 GAACCAACCTTACGGATGACTCGTGTCTTTGACCCCTTACACAGTGTCCCATTTATGCGTTGG 6480  
QY 6829 AGATGAGTGGGAACGAATCTCTGAATCAGGCTTTAAACTGTTGTGCGCAGTGTCTAGGCTT 6888  
DB 6481 AGATGAGTGGGAACGAATCTCTGAATCAGGCTTTAAACTGTTGTGCGCAGTGTCTAGGCTT 6540  
QY 6889 TGAAGTGTGTCATTTCCAGATGATTTCTATCTAGATGCTGCCATGACAAATGTTGTGACTA 6948  
DB 6541 TGAAGTGTGTCATTTCCAGATGATTTCTATCTAGATGCTGCCATGACAAATGTTGTGACTA 6600  
QY 6949 CAAGATTGGAGAGACGAGTGGACCGTCCAGGAGAAAATGCGCAGATGATGAGCTGCACATG 7008  
DB 6601 CAAGATTGGAGAGACGAGTGGACCGTCCAGGAGAAAATGCGCAGATGATGAGCTGCACATG 6660  
QY 7009 TCTTGGAAACGGAAGAGGAATTCAGTGTGACCCCTCATGAGGCAACCTGTTACGATGA 7068  
DB 6661 TCTTGGAAACGGAAGAGGAATTCAGTGTGACCCCTCATGAGGCAACCTGTTACGATGA 6720  
QY 7069 TGGGAGACATACCACTAGGAGAACAGTGGCAGAGAAATATCTCGGTGCCATTTGCTC 7128  
DB 6721 TGGGAGACATACCACTAGGAGAACAGTGGCAGAGAAATATCTCGGTGCCATTTGCTC 6780  
QY 7129 CTGCACATGCTTTTGGAGGCCAGCGGGCTGTGACAACTGTGACCGCAGACCTGCGGG 7188  
DB 6781 CTGCACATGCTTTTGGAGGCCAGCGGGCTGTGACAACTGTGACCGCAGACCTGCGGG 6840  
QY 7189 TGAACCCAGTCCCGAGAGGCACTACTGCGCAGTCTTACAAACAGTATTTCTCAGAGATACCA 7248  
DB 6841 TGAACCCAGTCCCGAGAGGCACTACTGCGCAGTCTTACAAACAGTATTTCTCAGAGATACCA 6900  
QY 7249 TCAGAGAAACAAACACTAAATGTTAAATTTGCCCAATTTGAGTGTCTTCATGCTTTAGATGACA 7308  
DB 6901 TCAGAGAAACAAACACTAAATGTTAAATTTGCCCAATTTGAGTGTCTTCATGCTTTAGATGACA 6960  
QY 7309 GGCTGACAGAGAGATTCGCGAGGTAAATCATCTTTCCAAATCCAGAGGAAACAGCATGT 7368  
DB 6961 GGCTGACAGAGAGATTCGCGAGGTAAATCATCTTTCCAAATCCAGAGGAAACAGCATGT 7020  
QY 7369 CTCTCTCCCAAGATCCATCTAAACTGGAGTGTAGTGTAGCAGACCCAGCTTAGAGTTCTTC 7428  
DB 7021 CTCTCTCCCAAGATCCATCTAAACTGGAGTGTAGTGTAGCAGACCCAGCTTAGAGTTCTTC 7080  
QY 7429 TTTCTTTTCTTAAGCCCTTGTCTGAGGAGTGTCTCCAGCTTCCAGCTCAGCTCAGCAT 7488  
DB 7081 TTTCTTTTCTTAAGCCCTTGTCTGAGGAGTGTCTGAGGTTTCTCATAAATGAGGCTGCACTT 7140  
QY 7489 TCTCAAGCATCACTTGGAGTTCCTGAGGTTTCTCATAAATGAGGCTGCACTT 7548  
DB 7141 TCTCAAGCATCACTTGGAGTTCCTGAGGTTTCTCATAAATGAGGCTGCACTT 7200  
QY 7549 GCCTGTTCTGCTGGAAGTATTCATACCGCTCAGTATTTTAAATGAGTGTACTAAGA 7608  
DB 7201 GCCTGTTCTGCTGGAAGTATTCATACCGCTCAGTATTTTAAATGAGTGTACTAAGA 7260  
QY 7609 TTTGTTTGGGATCAATAGGAAGCATATGACGCAACCAAGATGCAAAATGTTTGAAT 7668  
DB 7261 TTTGTTTGGGATCAATAGGAAGCATATGACGCAACCAAGATGCAAAATGTTTGAAT 7320  
QY 7669 GATATGACCAAAATTTTAAAGTAGGAAAGTCAACCAACACTTCTGCTTTCACTTAAGTGT 7728  
DB 7321 GATATGACCAAAATTTTAAAGTAGGAAAGTCAACCAACACTTCTGCTTTCACTTAAGTGT 7380  
QY 7729 CTGCGCCGCAATCTGAGGAGCAAGCATGATCTGTTACTGTGATATTTTAAATATCCA 7788  
DB 7381 CTGCGCCGCAATCTGAGGAGCAAGCATGATCTGTTACTGTGATATTTTAAATATCCA 7440

2y	7789	CAGTACT 7795	1009	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	1068
Db	7441	CAGTACT 7447	661	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
RESULT 12					
US-10-447-161-4					
; Sequence 4, Application US/10447.161					
; Publication No. US20040023314A1					
; GENERAL INFORMATION:					
; APPLICANT: Wang, Rong-fu					
; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis					
; FILE REFERENCE: HO-P0248US1					
; CURRENT APPLICATION NUMBER: US/10/447.161					
; CURRENT FILING DATE: 2003-05-28					
; PRIOR APPLICATION NUMBER: 60/383,530					
; PRIOR FILING DATE: 2002-05-28					
; NUMBER OF SEQ ID NOS: 148					
; SOFTWARE: PatentIn version 3.1					
; SEQ ID NO 4					
; LENGTH: 7705					
; TYPE: DNA					
; ORGANISM: Human					
US-10-447-161-4					
Query Match 95.2%; Score 7418.2; DB 16; Length 7705;					
Best Local Similarity 99.8%; Pred. No. 0;					
Matches 7429; Conservative 0; Mismatches 18; Indels 0; Gaps 0;					
Qy	349	GAAGAGCAAGAGGAGGAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTGGTGTGCTCA	408	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	1068
Db	1	GAAGAGCAAGAGGAGGAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTGGTGTGCTCA	60	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	409	AAGCAAGCCCGTTGTTATGATGCAATGGAACACATATCAGATTAATCAACAGTGGGAGCG	468	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	1068
Db	61	AAGCAAGCCCGTTGTTATGATGCAATGGAACACATATCAGATTAATCAACAGTGGGAGCG	120	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	469	GACCTACCTAGCGCAATGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	528	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	1068
Db	121	GACCTACCTAGCGCAATGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	180	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	529	CGAAGTAACTGAGCTGAGAGACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	588	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	1068
Db	181	CGAAGTAACTGAGCTGAGAGACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	240	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	589	AGTGGGTGACACTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACTGCATCGG	648	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	241	AGTGGGTGACACTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACTGCATCGG	300	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	649	GGTGGGCGAGGAGATTAAGCTGTACCATCGCAAAACCGCTGCCATGAGAGGGGTGATGC	708	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	301	GGTGGGCGAGGAGATTAAGCTGTACCATCGCAAAACCGCTGCCATGAGAGGGGTGATGC	360	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	709	CTACAAGATTGGTGACCTCGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG	768	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	361	CTACAAGATTGGTGACCTCGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG	420	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	769	TGTGTGTTTGGTAATGAAAGGAGAAATGCACTTGAAGCCCATAGCTGAGAGAGTGT	828	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	421	TGTGTGTTTGGTAATGAAAGGAGAAATGCACTTGAAGCCCATAGCTGAGAGAGTGT	480	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	829	TGATCATGCTGCTGGGACTTCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	888	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	481	TGATCATGCTGCTGGGACTTCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	540	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	889	CTGGATGATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	948	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	541	CTGGATGATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	600	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	949	TAGAAATAGATGCAACGATCAGGACACAGGACATCTTATAGATTGGAGACACCTGGAG	1008	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	601	TAGAAATAGATGCAACGATCAGGACACAGGACATCTTATAGATTGGAGACACCTGGAG	660	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1009	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	1068	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	1068
Db	661	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1069	GTGGAAGTGTGAGAGGACACCTCTGTGTCAGACCAATCGAGCGGATCTGGCCCTTCAC	1128	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	721	GTGGAAGTGTGAGAGGACACCTCTGTGTCAGACCAATCGAGCGGATCTGGCCCTTCAC	780	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1129	CGATGTTGTCAGAGTGTGTTTACCAACCCGAGCTCACCCGAGCTCTCTCTATGGCCA	1188	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	781	CGATGTTGTCAGAGTGTGTTTACCAACCCGAGCTCACCCGAGCTCTCTCTATGGCCA	840	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1189	CTGTGTCAAGACAGTGTGTGTTTACTCTGTGGGGATGAGTGGCTGGAAGACACAAGG	1248	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	841	CTGTGTCAAGACAGTGTGTGTTTACTCTGTGGGGATGAGTGGCTGGAAGACACAAGG	900	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1249	AAATAAGCAAAATGCTTTCAGCGTCTGGGCAACGAGTCTAGCTCCCAAGAGACAGCTGT	1308	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	901	AAATAAGCAAAATGCTTTCAGCGTCTGGGCAACGAGTCTAGCTCCCAAGAGACAGCTGT	960	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1309	AACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCATGTGTCTTACCATTCACTCAAA	1368	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	961	AACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCATGTGTCTTACCATTCACTCAAA	1020	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1369	TGCGAGGAGCTTACTCTCTGCAACCAAGAGGGCGAGGACGACATCTTGTGTGAG	1428	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1021	TGCGAGGAGCTTACTCTCTGCAACCAAGAGGGCGAGGACGACATCTTGTGTGAG	1080	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1429	CACAACTTCGAAATGAGCAGGACCAAGAAATCTCTTCTGCAAGACCACTCTTTT	1488	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1081	CACAACTTCGAAATGAGCAGGACCAAGAAATCTCTTCTGCAAGACCACTCTTTT	1140	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1489	GGTTGAGACTCGAGGAGGAAAATTCAAATGGTGCCTTGTGCCACTTCCCCTTCTTACAA	1548	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1141	GGTTGAGACTCGAGGAGGAAAATTCAAATGGTGCCTTGTGCCACTTCCCCTTCTTACAA	1200	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1549	CAACCAAAATACATGATTGCACTTCTGAGGGCAGAGAGACACATGAAAGTGTGTGG	1608	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1201	CAACCAAAATACATGATTGCACTTCTGAGGGCAGAGAGACACATGAAAGTGTGTGG	1260	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1609	GACCAACAGAACTATGATGCCGACCCAGAAATTTGGGTTTCTGCCCATGGCTGCCACGA	1668	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1261	GACCAACAGAACTATGATGCCGACCCAGAAATTTGGGTTTCTGCCCATGGCTGCCACGA	1320	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1669	GGAAATCTGCAACCAATGAAAGGGGTCTATGATCCGCAATGGAGATCAGTGGGATAGCA	1728	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1321	GGAAATCTGCAACCAATGAAAGGGGTCTATGATCCGCAATGGAGATCAGTGGGATAGCA	1380	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1729	GCATGACATGGGTCAATGATGAGTGCACGTGTGTTGGGAATGCTGCTGGGGAATGGAC	1788	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1381	GCATGACATGGGTCAATGATGAGTGCACGTGTGTTGGGAATGCTGCTGGGGAATGGAC	1440	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1789	ATGCAATGCTACTCGCAGCTTCCAGATCAGTGCATGTTGTTGATGACATCACTTCAATGT	1848	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1441	ATGCAATGCTACTCGCAGCTTCCAGATCAGTGCATGTTGTTGATGACATCACTTCAATGT	1500	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1849	GACGACACATTTCCCAAGCGTCATGAAGAGGGGACATGCTGAACTGTACATGCTTCGG	1908	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1501	GACGACACATTTCCCAAGCGTCATGAAGAGGGGACATGCTGAACTGTACATGCTTCGG	1560	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1909	TCAGGTCGCGGGCAGGTGGAAGTGTATCCCGTCGACCAATGCCAGGATTCAGAGACTGG	1968	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1561	TCAGGTCGCGGGCAGGTGGAAGTGTATCCCGTCGACCAATGCCAGGATTCAGAGACTGG	1620	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	1969	GAGCTTTTATCAAAATGGAGATTCATGAGGAGATGTTGTCATGTTGTCAGATACCAAGT	2028	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1621	GAGCTTTTATCAAAATGGAGATTCATGAGGAGATGTTGTCATGTTGTCAGATACCAAGT	1680	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	2029	CTACTGCTATGGCCGTCGTCATGGGAGTGGCAATGCCAACCTTTTACAGACCTATCCAAG	2088	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Db	1681	CTACTGCTATGGCCGTCGTCATGGGAGTGGCAATGCCAACCTTTTACAGACCTATCCAAG	1740	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720
Qy	2089	CTCAAGTGTCTGTGCAAGATTTTATCACTGAGACTCGAGTCCGAGCCCACTCCCAACC	2148	CAAGAAAGGATATCGAGGAAAACCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGA	720



4309 CACACCTCTTTAGAGGAGACAGAAAACAGGTCTTGTATTCGCCAACTGGGATTGACTTTTC 4368  
4361 CACACCTCTTTAGAGGAGACAGAAAACAGGTCTTGTATTCGCCAACTGGGATTGACTTTTC 4020  
4369 TGATATTACTGCCAACTCTTTTACTGTGACTGGATTGCTCTCGAGCCACCATCACTGG 4428  
4021 TGATATTACTGCCAACTCTTTTACTGTGCACTGGATTGCTCTCGAGCCACCATCACTGG 4080  
4429 CTACAGGATCCGCCATCATCCGAGCACTTCAGTGGGAGACCTCGAGAAAGATCGGGTGCC 4488  
4081 CTACAGGATCCGCCATCATCCGAGCACTTCAGTGGGAGACCTCGAGAAAGATCGGGTGCC 4140  
4489 CCACCTCTCGGAATTCATACCCCTCACCACCTCACTCCAGGACACAGATATGTGTGTCAG 4548  
4141 CCACCTCTCGGAATTCATACCCCTCACCACCTCACTCCAGGACACAGATATGTGTGTCAG 4200  
4549 CATCGTTGCTCTTAATGGCAGAGAGGAAGTCCCTTATTGATTGGGCCAAACAATCAACAGT 4608  
4201 CATCGTTGCTCTTAATGGCAGAGAGGAAGTCCCTTATTGATTGGGCCAAACAATCAACAGT 4260  
4609 TTCTGATGTTCCGAGGACCTCGGAAGTTGTTGTCGACCCCAACAGCCTCACTGATCAG 4668  
4261 TTCTGATGTTCCGAGGACCTCGGAAGTTGTTGTCGACCCCAACAGCCTCACTGATCAG 4320  
4669 CTGGGATGCTCTGCTCTCACTGAGTATATACAGGATCACTTACGAGAAACAGGAGG 4728  
4321 CTGGGATGCTCTGCTCTCACTGAGTATATACAGGATCACTTACGAGAAACAGGAGG 4380  
4729 AAATAGCCCTGTCAGAGGATCACTGCTGGGAGCAAGTCTACAGTACCATCAGCGG 4788  
4381 AAATAGCCCTGTCAGAGGATCACTGCTGGGAGCAAGTCTACAGTACCATCAGCGG 4440  
4789 CCTTAAACCTGGAGTATATACCATCACTGTATGCTGCTCACTGGCGTGGAGACAG 4848  
4441 CCTTAAACCTGGAGTATATACCATCACTGTATGCTGCTCACTGGCGTGGAGACAG 4500  
4849 CCCGCAAGCAGCAAGCAATTTCCATTAATACCGAAACAGAAATGACAAAACATCCCA 4908  
4501 CCCGCAAGCAGCAAGCAATTTCCATTAATACCGAAACAGAAATGACAAAACATCCCA 4560  
4909 GATCGAATGACGATGTTACGACACACAGTATGCTCAAGTGGCTGCTTCAAGTTC 4968  
4561 GATCGAATGACGATGTTACGACACACAGTATGCTCAAGTGGCTGCTTCAAGTTC 4620  
4969 CCCTGTTTACTGTTTACAGAGTAAACCACTCTCCAAAATAGGACCAAGCAACAAAAC 5028  
4621 CCCTGTTTACTGTTTACAGAGTAAACCACTCTCCAAAATAGGACCAAGCAACAAAAC 4680  
5029 TAAACCTGAGGTCAGATCAACAGAAATGACTATTGAAGGCTTCAGCCCAAGTGA 5088  
4681 TAAACCTGAGGTCAGATCAACAGAAATGACTATTGAAGGCTTCAGCCCAAGTGA 4740  
5089 GTATGTTGTTAGTGTCTATGCTCAGAAATCAAGGGGAGAGTCAAGCTCTGTTGTCAGAC 5148  
4741 GTATGTTGTTAGTGTCTATGCTCAGAAATCAAGGGGAGAGTCAAGCTCTGTTGTCAGAC 4800  
5149 TGCAGTAAACCAATGATGATGCCCTTAAGGACTGGCAATTCATGATGTGATGTGCAATTC 5208  
4801 TGCAGTAAACCAATGATGATGCCCTTAAGGACTGGCAATTCATGATGTGATGTGCAATTC 4860  
5209 CATCAAAATTCCTTGGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGCTGACTTACTC 5268  
4861 CATCAAAATTCCTTGGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGCTGACTTACTC 4920  
5269 GAGCCCTGAGGATGGAATTCATAGCTATTTCCTGCACTGATGTTGGAAGAGACTGTC 5328  
4921 GAGCCCTGAGGATGGAATTCATAGCTATTTCCTGCACTGATGTTGGAAGAGACTGTC 4980  
5329 AGAGCTCAAGGCTTCAGCCGGGTTCTGAGTACACAGTCACTGTTGCTTGCACGA 5388  
4981 AGAGCTCAAGGCTTCAGCCGGGTTCTGAGTACACAGTCACTGTTGCTTGCACGA 5040

5389 TGATATGAGAGCCAGGCCCTGATTGGAAACCACTCCACAGCTATTCTCTGCACCAACTGA 5448  
5041 TGATATGAGAGCCAGGCCCTGATTGGAAACCACTCCACAGCTATTCTCTGCACCAACTGA 5100  
5449 CCTGAAGTTCACTCAGGTCAACCAAGCTCAGCGCCAGTGGACACCAACCAATGT 5508  
5101 CCTGAAGTTCACTCAGGTCAACCAAGCTCAGCGCCAGTGGACACCAACCAATGT 5160  
5509 TCAGCTCACTGGATATCGAGTGCAGTGCACCCCAAGGAGAAAGACCGGACCAATGAAGA 5568  
5161 TCAGCTCACTGGATATCGAGTGCAGTGCACCCCAAGGAGAAAGACCGGACCAATGAAGA 5220  
5569 AATCAACCTTCTCTGACAGCTCATCGTGTGTTGATCAGAGCTTATGTTGGGCCACCAA 5628  
5221 AATCAACCTTCTCTGACAGCTCATCGTGTGTTGATCAGAGCTTATGTTGGGCCACCAA 5280  
5629 ATATGAAGTGTGATCTATGCTCTTAAGGACACTTTGACAGCAGACGAGCTCAGGGTGT 5688  
5281 ATATGAAGTGTGATCTATGCTCTTAAGGACACTTTGACAGCAGACGAGCTCAGGGTGT 5340  
5689 TGTCAACCTCTGAGAGATGTCAGCCCAAGAGGGCTCGTGTGACAGATGCTACTGA 5748  
5341 TGTCAACCTCTGAGAGATGTCAGCCCAAGAGGGCTCGTGTGACAGATGCTACTGA 5400  
5749 GACCAACCTATCACTTAGCTGGAGAACCAAGACTTGAGACGATCACTGGCCTTCCAAAGTGA 5808  
5401 GACCAACCTATCACTTAGCTGGAGAACCAAGACTTGAGACGATCACTGGCCTTCCAAAGTGA 5460  
5809 TGCCCTTCCAGCAATGAGCAGATCCCAATCCAGAGACCAATCAAGCAGATGTCAGAG 5868  
5461 TGCCCTTCCAGCAATGAGCAGATCCCAATCCAGAGACCAATCAAGCAGATGTCAGAG 5520  
5869 CTACACCACTACTGCTTACACCAAGGCACTGACTTACAGATCTACCTGTACACCTTGAA 5928  
5521 CTACACCACTACTGCTTACACCAAGGCACTGACTTACAGATCTACCTGTACACCTTGAA 5580  
5929 TGACAACTCTGAGCTCCCTGTGTGTCATCGACGCTTCCACTGCGCATGATGACCAATC 5988  
5581 TGACAACTCTGAGCTCCCTGTGTGTCATCGACGCTTCCACTGCGCATGATGACCAATC 5640  
5989 CAACCTGCTGCTTCCGCGCACCAACCAATTCCTTGTGTATCATGCGACGCGCCACG 6048  
5641 CAACCTGCTGCTTCCGCGCACCAACCAATTCCTTGTGTATCATGCGACGCGCCACG 5700  
6049 TGCCAGGATACCGGCTACATCATCAAGTATGAGAGCTGGGTCTCTCCAGAGAAAT 6108  
5701 TGCCAGGATACCGGCTACATCATCAAGTATGAGAGCTGGGTCTCTCCAGAGAAAT 5760  
6109 GGTCCCTCGGCCCGCTGTGTGTCACAGAGGCTACTATTACTGGCCTGGAACCGGAAAC 6168  
5761 GGTCCCTCGGCCCGCTGTGTGTCACAGAGGCTACTATTACTGGCCTGGAACCGGAAAC 5820  
6169 CGAATATCAATTTATGTCAATTGCTGCTGAAGAAATATCAGAAAGACGAGCCCTGATTGG 6228  
5821 CGAATATCAATTTATGTCAATTGCTGCTGAAGAAATATCAGAAAGACGAGCCCTGATTGG 5880  
6229 AAGGAAAGACAGAGAGCTTCCCACTGTAACCTTCCACACCCCAATCTTCATGG 6288  
5881 AAGGAAAGACAGAGAGCTTCCCACTGTAACCTTCCACACCCCAATCTTCATGG 5940  
6289 ACCAGAGATCTTGGATGTTCTCTCCACAGTTTCAAAAGACCCCTTTCGTACCCACCCCTGG 6348  
5941 ACCAGAGATCTTGGATGTTCTCTCCACAGTTTCAAAAGACCCCTTTCGTACCCACCCCTGG 6000  
6349 GTATCACACTGGAATGATTTTCACTTCTCGGCACTTCTGTCAGCAACCCAGTGTGG 6408  
6001 GTATCACACTGGAATGATTTTCACTTCTCGGCACTTCTGTCAGCAACCCAGTGTGG 6060  
6409 GCAACAAATGATCTTTTGGAGAACTGTTTTAGGCGGACCAACCGCCCAACACCGGCCAC 6468  
6061 GCAACAAATGATCTTTTGGAGAACTGTTTTAGGCGGACCAACCGCCCAACACCGGCCAC 6120  
6469 CCCCATAGGCAATAGCCCAAGCAATATCCCGCCGATGTAGGACAAAGACTCTCTCTCA 6528



1b		CCCCATTAAGGCATAGGCCAAGACCAATACCCGCCGAATGTAGGACAAAGACTCTCTCTCA	6180
1y		GACAACCATCTCATGGGCCCCATTCAGGACACACTTCTGAGTACATCATTTTCATGTCATCC	6598
1b		GACAACCATCTCATGGGCCCCATTCAGGACACACTTCTGAGTACATCATTTTCATGTCATCC	6240
1y		TGTTGGCACTGATGTAAGAACCTTACAGTTTCAGGGTTCCTCGAACTTCTACCAAGTCCAC	6648
1b		TGTTGGCACTGATGTAAGAACCTTACAGTTTCAGGGTTCCTCGAACTTCTACCAAGTCCAC	6300
1y		TCTGACAGGCTTCAACAGAGGTGCCACCTACAACATCATAGTGAGGACCTGAAGAACA	6708
1b		TCTGACAGGCTTCAACAGAGGTGCCACCTACAACATCATAGTGAGGACCTGAAGAACA	6360
1y		GCAGAGCATTAAGTTTCGGGAAGAGTGTCTACCTGGGCAACTCTCTCAACAGAGCTT	6768
1b		GCAGAGCATTAAGTTTCGGGAAGAGTGTCTACCTGGGCAACTCTCTCAACAGAGCTT	6420
1y		GAACCAACTACGGATGACTCGTGTCTTGACCCCTACACAGTGTCCCATTTATGCCGTGG	6828
1b		GAACCAACTACGGATGACTCGTGTCTTGACCCCTACACAGTGTCCCATTTATGCCGTGG	6480
1y		AGATGAGTGGGAACGAATGCTGCAATCAGGCTTTAAACGTGTTGGCCAGTGTCTTAGGCTT	6888
1b		AGATGAGTGGGAACGAATGCTGCAATCAGGCTTTAAACGTGTTGGCCAGTGTCTTAGGCTT	6540
1y		TGGAAGTGGTCAATTCAGATGTGATTCATCTATAGTGGTGCATGACAAATGGTGTGAACATA	6948
1b		TGGAAGTGGTCAATTCAGATGTGATTCATCTATAGTGGTGCATGACAAATGGTGTGAACATA	6600
1y		CAGATTGGAGAGAGTGGGACCGTCAGGAGAAAATGGCCAGATGATGAGCTGCACATG	7008
1b		CAGATTGGAGAGAGTGGGACCGTCAGGAGAAAATGGCCAGATGATGAGCTGCACATG	6660
1y		TCTTGGAAACGGAAAAGGAAATTCAGTGTGACCCCTCATGAGGCAACGTGTTACGATGA	7068
1b		TCTTGGAAACGGAAAAGGAAATTCAGTGTGACCCCTCATGAGGCAACGTGTTACGATGA	6720
1y		TGGGAAGACATACACGTAGGAGAACAGTGGCAGAAAGAAATCTCGGTGCAATTTGCTC	7128
1b		TGGGAAGACATACCAAGTAGGAGAACAGTGGCAGAAAGAAATCTCGGTGCAATTTGCTC	6780
1y		CTGCACATGCTTTGGAGGCAGCGGGCTGGCGCTGTGACAACTGCGCGACACTGGGGG	7188
1b		CTGCACATGCTTTGGAGGCAGCGGGCTGGCGCTGTGACAACTGCGCGACACTGGGGG	6840
1y		TGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTTACAAACAGATATCTCAGAGTACCA	7248
1b		TGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTTACAAACAGATATCTCAGAGTACCA	6900
1y		TCAGAGAACAAACATAATGTTAAATGGCCAAATTCAGTGTTCATGCTTTAGAGTACCA	7308
1b		TCAGAGAACAAACATAATGTTAAATGGCCAAATTCAGTGTTCATGCTTTAGAGTACCA	6960
1y		GGCTGACAGAGAGATTCGCCAGAGTAAATCATCTTCCATTCAGAGAACAAAGCATGT	7368
1b		GGCTGACAGAGAGATTCGCCAGAGTAAATCATCTTCCATTCAGAGAACAAAGCATGT	7020
1y		CTCTCTGCAAGATCCATCTAAACTGGAGTGTAGCAGACCCAGCTTAGAGTCTTCTC	7428
1b		CTCTCTGCAAGATCCATCTAAACTGGAGTGTAGCAGACCCAGCTTAGAGTCTTCTC	7080
1y		TTTCTTTCTTAAAGCCCTTGCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCACAGCT	7488
1b		TTTCTTTCTTAAAGCCCTTGCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCACAGCT	7140
1y		TCTCGAAGCATCACCTGGAGGTTTTCTGAGGGTTTTCTCATAAATGAGGGCTGCACATT	7548
1b		TCTCGAAGCATCACCTGGAGGTTTTCTGAGGGTTTTCTCATAAATGAGGGCTGCACATT	7200
1y		GCCTGTTCTGCTTCGAAAGTATTCAAATACCGCTCAGTATTTTAAATGAAGTCAATTCGAAGA	7608

Db	7201	GCCTGTTCTGCTCGAAGTATTTC	CAATACCGCTCAGTATTATTTAAATGAAGTGATTTCTAAGA	72668
Qy	7609	TTTGGTTTGGGATCAATAGGAAGAC	NATGCGCCCAACCAAGATGCAAAATGTTTGAAT	7668
Db	7261	TTTGGTTTGGGATCAATAGGAAGAC	NATGCGCCCAACCAAGATGCAAAATGTTTGAAT	7320
Qy	7669	GATATGACCAAAATTTTAAAGTAG	AGAAAGTCACCCAAACACTTCTGCTTTTCACTTAAAGTGT	7728
Db	7321	GATATGACCAAAATTTTAAAGTAG	AGAAAGTCACCCAAACACTTCTGCTTTTCACTTAAAGTGT	7380
Qy	7729	CTGGCCCGCAATACTGTATGGAAC	AGCATGATCTTGTACTGTGATATTTAAATATCCA	7788
Db	7381	CTGGCCCGCAATACTGTATGGAAC	AGCATGATCTTGTACTGTGATATTTAAATATCCA	7440
Qy	7789	CAGTACT	7795	
Db	7441	CAGTACT	7447	

RESULT 13  
 US-10-098-841-6  
 ; Sequence 6, Application US/10098841  
 ; Publication No. US20020197679A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Liu, Chenghua  
 ; APPLICANT: Asundi, Vinod  
 ; APPLICANT: Xu, Chongjun  
 ; APPLICANT: Zhou, Ping  
 ; APPLICANT: Ma, Yungling  
 ; APPLICANT: Wang, Jian-Rui  
 ; APPLICANT: Zhao, Qing Ai  
 ; APPLICANT: Ren, Feiyan  
 ; APPLICANT: Chen, Rui-hong  
 ; APPLICANT: Wang, Dunrui  
 ; APPLICANT: Wang, Zhiwei  
 ; APPLICANT: Wehrman, Tom  
 ; APPLICANT: Zhang, Jie  
 ; APPLICANT: Qian, Xiaohong B.  
 ; APPLICANT: Drmanac, Radoje T.  
 ; TITLE OF INVENTION: No. US20020197679A1elel Nucleic Acids and  
 ; TITLE OF INVENTION: Polypeptides  
 ; FILE REFERENCE: 784CIP2  
 ; CURRENT APPLICATION NUMBER: US/10/098,841  
 ; CURRENT FILING DATE: 2002-03-13  
 ; PRIOR APPLICATION NUMBER: 09/598,042  
 ; PRIOR FILING DATE: 2000-06-20  
 ; PRIOR APPLICATION NUMBER: 09/552,317  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: 09/488,725  
 ; PRIOR FILING DATE: 2000-01-21  
 ; NUMBER OF SEQ ID NOS: 331  
 ; SOFTWARE: pt\_Fl\_genes Version 1.0  
 ; SEQ ID NO 6  
 ; LENGTH: 7867  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (363)..(7160)  
 US-10-098-841-6

Qy	10	CCGCTGTGTCTGCACAGGGGGAGGAGGGAA	CCCCCAGGGCGCAGCGGGGAAGAGGGGACC	69
Db	104	CCGCTGTGTCTGCACAGGGGGAGGAGGGAA	CCCCCAGGGCGCAGCGGGGAAGAGGGGACC	163
Qy	70	TGCAGCCACAACTTCTCTGTCTCTGCA	ATCCCTTCTGTCCCTCCACCCCGTCCCTTCCC	129
Db	164	TGAGCCACAACTTCTCTGTCTCTGCA	ATCCCTTCTGTCCCTCCACCCCGTCCCTTCCC	223

Query Match 92.7%; Score 7224.8; DB 13; Length 7867;  
 Best local Similarity 96.4%; Pred. No. 0;  
 Matches 7509; Conservative 0; Mismatches 7; Indels 270; Gaps 1;

	Query_Match	92.7%	Score 2224.8	DB 13	Length 7867
	Best_Local_Similarity	96.4%	Pred. Nc. 0		
	Matches 7509	Conservative	0	Mismatches 7	Indels 270
					Gaps 1
10	CCGCGTGTGCTGCAACAGGGGAGAGAGGGAA	CCCCAGGCGCGAGCGGGGAAGAGGGGACC	69		
104	CCGCGTGTGCTGCAACAGGGGAGAGAGGGAA	CCCCAGGCGCGAGCGGGGAAGAGGGGACC	163		
70	TGCAGCCACAACTTCTCTGTGCTCTCATCCCTT	CTGTCTCCCTCCACCCCGTCCCTCTCCC	129		
164	TGCAGCCACAACTTCTCTGTGCTCTCATCCCTT	CTGTCTCCCTCCACCCCGTCCCTCTCCC	223		



QY 130 CACCTCTGGCCGCCACCTTCTTGGAGCGCAAAACCCCGGGAGGCATTAAGAGGATTT 189  
 Db 224 CACCTCTGGCCGCCACCTTCTTGGAGCGCAAAACCCCGGGAGGCATTAAGAGGATTT 283  
 QY 190 TTCGCGCAGGTTCGAAAGGGAAGCAAACTTGGTGGCAACTTGCCTCCGCGTGGCGGCTC 249  
 Db 284 TTCGCGCAGGTTCGAAAGGGAAGCAAACTTGGTGGCAAACTTGCCTCCGCGTGGCGGCTC 343  
 QY 250 TCTCCGCCACCGCTCTCAACATGCTTATAGGGGTTCGGGGGCCCGGGTCTGCTGCTGGCGGT 309  
 Db 344 TCTCCGCCACCGCTCTCAACATGCTTATAGGGGTTCGGGGGCCCGGGTCTGCTGCTGGCGGT 403  
 QY 310 CCTGTCCTGGGACAGCGGTGCTCCACGCGAGCCTTCGAAGAGCAAGAGCGAGGCTCA 369  
 Db 404 CCTGTCCTGGGACAGCGGTGCTCCACGCGAGCCTTCGAAGAGCAAGAGCGAGGCTCA 463  
 QY 370 GCAATGTTGTCAGCCCGCAGTCCCGGTGCTCAGTCAAAAGCAAGCCCGGTGCTTATGA 429  
 Db 464 GCAATGTTGTCAGCCCGCAGTCCCGGTGCTCAGTCAAAAGCAAGCCCGGTGCTTATGA 523  
 QY 430 CAATGGAAACACTATCAGATAAAATCAACAGTGGAGCGGACCTACCTAGGCAATGGGTT 489  
 Db 524 CAATGGAAACACTATCAGATAAAATCAACAGTGGAGCGGACCTACCTAGGCAATGGGTT 583  
 QY 490 GGTTTGTACTTGTATGGAGGAGCGGAGGTTTAACTGCGAAAGTAAACCTGAAGCTGA 549  
 Db 584 GGTTTGTACTTGTATGGAGGAGCGGAGGTTTAACTGCGAAAGTAAACCTGAAGCTGA 643  
 QY 550 AGAGACTTGTCTTGAACAAGTACACTGGGAACACTTACCGAGTGGGTGACACTTATGAGCG 609  
 Db 644 AGAGACTTGTCTTGAACAAGTACACTGGGAACACTTACCGAGTGGGTGACACTTATGAGCG 703  
 QY 610 TCCTAAAGCTCCATGATCTGGGACTGTACCTGCTATCGGAGTGGGCGGAGGAGGATAAG 669  
 Db 704 TCCTAAAGCTCCATGATCTGGGACTGTACCTGCTATCGGAGTGGGCGGAGGAGGATAAG 763  
 QY 670 CTGTACCATCGCAAAACCGCTGCATGAAAGGGGGTCACTTACAGATTTGGTGACACCTG 729  
 Db 764 CTGTACCATCGCAAAACCGCTGCATGAAAGGGGGTCACTTACAGATTTGGTGACACCTG 823  
 QY 730 GAGGAGACACATGAGACTGTGGTGGTTACATGTTAGAGTGTGTGTCTTGTGTAATGGAAA 789  
 Db 824 GAGGAGACACATGAGACTGTGGTGGTTACATGTTAGAGTGTGTGTCTTGTGTAATGGAAA 883  
 QY 790 AGGAGAACTGGACCTGCAAGCCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTC 849  
 Db 884 AGGAGAACTGGACCTGCAAGCCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTC 943  
 QY 850 CTATGTTGTCGAGAAAACGTGGGAGAACCTTACCAAGCTGGATGATGGTAGATTGTAC 909  
 Db 944 CTATGTTGTCGAGAAAACGTGGGAGAACCTTACCAAGCTGGATGATGGTAGATTGTAC 1003  
 QY 910 TTGCTCTGGGAGAGGAGCGGAGCGCATCACTTGGCACTTCTAGAAATAGATGCAAGATCA 969  
 Db 1004 TTGCTCTGGGAGAGGAGCGGAGCGCATCACTTGGCACTTCTAGAAATAGATGCAAGATCA 1063  
 QY 970 GGAACAAGGACATCTATAGAAATTGGAGACACCTGGAGCAAGAGGATAATTCGAGGAAA 1029  
 Db 1064 GGAACAAGGACATCTATAGAAATTGGAGACACCTGGAGCAAGAGGATAATTCGAGGAAA 1123  
 QY 1030 CTTGCTCCAGTGCATCTGCACAGGCAACCGCGAGGAGTGGAGTGTGAGAGGCACAC 1089  
 Db 1124 CTTGCTCCAGTGCATCTGCACAGGCAACCGCGAGGAGTGGAGTGTGAGAGGCACAC 1183  
 QY 1090 CTTGCTGCAGACACATCGAGCGGATCTGGCCCTTACCGATGTTGCTGTCAGCTGTTTA 1149  
 Db 1184 CTTGCTGCAGACACATCGAGCGGATCTGGCCCTTACCGATGTTGCTGTCAGCTGTTTA 1243  
 QY 1150 CCNACCGGACCTCACCCCGAGCTCTCCCTTATGGCCACTGTGTCAAGACAGTGGTGT 1209  
 Db 1244 CCNACCGGACCTCACCCCGAGCTCTCCCTTATGGCCACTGTGTCAAGACAGTGGTGT 1303

QY 1210 GGTCTACTCTGTGGGATGCACTGCTGAAGACACAAGGAAATAAGAAATGCTTTTGCA 1269  
 Db 1304 GGTCTACTCTGTGGGATGCACTGCTGAAGACACAAGGAAATAAGAAATGCTTTTGCA 1363  
 QY 1270 GTGCTCTGGGCAACGGAGTCAGCTGCCAAGAGACAGTGTAAACCCAGACTTACGGTGGCAA 1329  
 Db 1364 GTGCTCTGGGCAACGGAGTCAGCTGCCAAGAGACAGTGTAAACCCAGACTTACGGTGGCAA 1423  
 QY 1330 CTCAATGGAGAGCCATGCTTACCACTTCACTTACCAATGGCAGGAGCTTACTCTCTG 1389  
 Db 1424 CTCAATGGAGAGCCATGCTTACCACTTCACTTACCAATGGCAGGAGCTTACTCTCTG 1483  
 QY 1390 CACCAAGAGGGCGACAGAGCGGACATCTTGGTGGACACAACCTTGAATTTAGACA 1449  
 Db 1484 CACCAAGAGGGCGACAGAGCGGACATCTTGGTGGACACAACCTTGAATTTAGACA 1543  
 QY 1450 GGAACCAAGAAATCTCTTCTGCAAGACCACTGTTTGGTTCAGACTTCGAGGAGGAAA 1509  
 Db 1544 GGAACCAAGAAATCTCTTCTGCAAGACCACTGTTTGGTTCAGACTTCGAGGAGGAAA 1603  
 QY 1510 TTCCAATGGTGGCTTGTGCACTTCCCTTCTATACACACCACTTACACTGATG 1569  
 Db 1604 TTCCAATGGTGGCTTGTGCACTTCCCTTCTATACACACCACTTACACTGATG 1663  
 QY 1570 CACTTCTGAGGGCAGAGAGACAACATGAAGTGGTGGTGGGACCAACAGAACTATGATGC 1629  
 Db 1664 CACTTCTGAGGGCAGAGAGACAACATGAAGTGGTGGTGGGACCAACAGAACTATGATGC 1723  
 QY 1630 GACACCAAGATTTGGGTTCTGCCCATGCTGCCACGAGGAAATCTGCACACCACTGA 1689  
 Db 1724 GACACCAAGATTTGGGTTCTGCCCATGCTGCCACGAGGAAATCTGCACACCACTGA 1783  
 QY 1690 AGGGGTCACTGATCCGCAATTTGGAGATCAGTGGGATGAAGCAGCATGACATGGGTCACTGAT 1749  
 Db 1784 AGGGGTCACTGATCCGCAATTTGGAGATCAGTGGGATGAAGCAGCATGACATGGGTCACTGAT 1843  
 QY 1750 GAGGTGACGTGTGTTGGAAATGGTCTGGGGAAATGGACATGCAATGGCTACTCGCAGCT 1809  
 Db 1844 GAGGTGACGTGTGTTGGAAATGGTCTGGGGAAATGGACATGCAATGGCTACTCGCAGCT 1903  
 QY 1810 TCGAGATCAGTGCATTTGATGACATCACTTACAATGTGAACGACACATTCACAAAGCG 1869  
 Db 1904 TCGAGATCAGTGCATTTGATGACATCACTTACAATGTGAACGACACATTCACAAAGCG 1963  
 QY 1870 TCATGAAGGGGCAACATGCTGAACCTGTACATGCTTGGTTCAGGGTTCGGGCGAGGTGGA 1929  
 Db 1964 TCATGAAGGGGCAACATGCTGAACCTGTACATGCTTGGTTCAGGGTTCGGGCGAGGTGGA 2023  
 QY 1930 GTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGGGACGTTTTATCAAAATTCGAGA 1989  
 Db 2024 GTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGGGACGTTTTATCAAAATTCGAGA 2083  
 QY 1990 TTCAATGGGAGAAATGATGTCATGGTTCAGATACCAAGTCTACTGCTATGGCCGTGGCAT 2049  
 Db 2084 TTCAATGGGAGAAATGATGTCATGGTTCAGATACCAAGTCTACTGCTATGGCCGTGGCAT 2143  
 QY 2050 TGGGAGTGCATTTGCCAACCCTTACAGACCTTACAGACCTTACAAAGTCAAGTGCCTGCTGCAAGT 2109  
 Db 2144 TGGGAGTGCATTTGCCAACCCTTACAGACCTTACAGACCTTACAAAGTCAAGTGCCTGCTGCAAGT 2203  
 QY 2110 ATTATCACTGAGACTCCGAGTCCAGGACCACTCCACCCCATCCAGTGGAAATGCACACA 2169  
 Db 2204 ATTATCACTGAGACTCCGAGTCCAGGACCACTCCACCCCATCCAGTGGAAATGCACACA 2263  
 QY 2170 GCAATCTCAATTTCCAAAGTACATCTCAGGTGGAGACCTTAAATTTCTGTAGGCCGTTG 2229  
 Db 2264 GCAATCTCAATTTCCAAAGTACATCTCAGGTGGAGACCTTAAATTTCTGTAGGCCGTTG 2323  
 QY 2230 GAGGAAAGCTTACCATACAGGCCACCTTAAACTCCTTACACCATCAAAAGGCTGAAGCTGG 2289  
 Db 2324 GAGGAAAGCTTACCATACAGGCCACCTTAAACTCCTTACACCATCAAAAGGCTGAAGCTGG 2383  
 QY 2290 TGTGGTATACAGGGGCGAGCTCATCAGCATCCAGCATGAGCGGCCACCAAGAGTGAAGTGC 2349



QY 4510 CCTCACCACCTCACTCCAGGACAGAGTATGGTTCAGCATCGTTCCTTAATGCGAG 4569  
DB |||||  
QY 4604 CCTCACCACCTCACTCCAGGACAGAGTATGGTTCAGCATCGTTCCTTAATGCGAG 4663  
DB |||||  
QY 4570 AGAGGAAAGTCCCTTATTGATTGGCCAAACAATCAACAGTTTCTGATGTTCCGAGGGACCT 4629  
DB |||||  
QY 4664 AGAGGAAAGTCCCTTATTGATTGGCCAAACAATCAACAGTTTCTGATGTTCCGAGGGACCT 4723  
DB |||||  
QY 4630 GGAAGTTTGGTGGCGGCCCCACACAGCTTACTGATGAGTGGGATGCTCTGCTGTCCAC 4689  
DB |||||  
QY 4724 GGAAGTTTGGTGGCGGCCCCACACAGCTTACTGATCAGCTGGGATGCTCTGCTGTCCAC 4783  
DB |||||  
QY 4690 AGTGAGATATTACAGGATCACTTACGGGAAACAGGAGGAAATAGCCCTGTCCAGGAGTT 4749  
DB |||||  
QY 4784 AGTGAGATATTACAGGATCACTTACGGGAAACAGGAGGAAATAGCCCTGTCCAGGAGTT 4843  
DB |||||  
QY 4750 CACTGTGGCTGGGACAGTCTACAGCTACCAATCAGCGGCTTAAACCTGAGTTGATTA 4809  
DB |||||  
QY 4844 CACTGTGGCTGGGACAGTCTACAGCTACCAATCAGCGGCTTAAACCTGAGTTGATTA 4903  
DB |||||  
QY 4810 TACCATCACTGTGTATGCTCACTGGCGGTGGAGACAGCGCCCGCAAGCAGCAAGCCAAAT 4869  
DB |||||  
QY 4904 TACCATCACTGTGTATGCTCACTGGCGGTGGAGACAGCGCCCGCAAGCAGCAAGCCAAAT 4963  
DB |||||  
QY 4870 TTCCATTAAATACCGAAACAGAAATTTGA CAACCAATCCAGATGCCAGTGCAGATGTTCA 4929  
DB |||||  
QY 4964 TTCCATTAAATACCGAAACAGAAATTTGA CAACCAATCCAGATGCCAGTGCAGATGTTCA 5023  
DB |||||  
QY 4930 GGACAAACAGCATTAGTGTCAGTGCTGCTTCAAGTTCCCTGTACTGGTTACAGAGT 4989  
DB |||||  
QY 5024 GGACAAACAGCATTAGTGTCAGTGCTGCTTCAAGTTCCCTGTACTGGTTACAGAGT 5083  
DB |||||  
QY 4990 AACCAACCACTCCCAAAAATGACACAGGACCAACAAAACTAAAACTGCAGTCCAGATCA 5049  
DB |||||  
QY 5084 AACCAACCACTCCCAAAAATGACACAGGACCAACAAAACTAAAACTGCAGTCCAGATCA 5143  
DB |||||  
QY 5050 AACCAAAATGACTATTGAAGGCTGCAGCCACAGTGAGTATGTTAGTGTCTATGC 5109  
DB |||||  
QY 5144 AACCAAAATGACTATTGAAGGCTGCAGCCACAGTGAGTATGTTAGTGTCTATGC 5203  
DB |||||  
QY 5110 TCAGAAATCCAAAGCGGAGAGTACGCTCTGGTTCCAGTCCAGTAAACCAATTTGATCG 5169  
DB |||||  
QY 5204 TCAGAAATCCAAAGCGGAGAGTACGCTCTGGTTCCAGTCCAGTAAACCAATTTGATCG 5253  
DB |||||  
QY 5170 CCTTAAAGGACTGGCATTTCACTGATGTGGATGTCCATCAAAATTTGCTTGGGAAG 5229  
DB |||||  
QY 5254 ----- 5253  
DB |||||  
QY 5230 CCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTTCTCGAGCCCTGAGGATGGAATCCA 5289  
DB |||||  
QY 5254 ----- 5253  
DB |||||  
QY 5290 TGAGCTATTCCCTGCACCTGATGGTGAAGAAGACACTGCAGAGCTGCAGGCTCAGACC 5349  
DB |||||  
QY 5254 ----- 5253  
DB |||||  
QY 5350 GGGTTCTGAGTACACAGTCAAGTGTGGTTGGCTTGCATGATGATATGGAGAGCCAGCCCT 5409  
DB |||||  
QY 5254 ----- 5253  
DB |||||  
QY 5410 GATTGGAAACCCAGTCCACAGTATTCTGCAACCAACTGACCTGAAGTTCACTCAGGTCCAC 5469  
DB |||||  
QY 5254 -----CTATTCTGCAACCAACTGACCTGAAGTTCACTCAGGTCCAC 5293  
DB |||||  
QY 5470 ACCCAAGACCTTGAGCGCCCGAGTGACACCACTGTTTCCAGTCTCACTGGATATCGAGT 5529  
DB |||||  
QY 5294 ACCCAAGACCTTGAGCGCCCGAGTGACACCACTGTTTCCAGTCTCACTGGATATCGAGT 5353  
DB |||||  
QY 5530 GCGGTGACCCCAAGGAGAAACCGGACCAATGAAGAAATCAACCTTGTCTCTGACAG 5589  
DB |||||  
QY 5354 GCGGTGACCCCAAGGAGAAACCGGACCAATGAAGAAATCAACCTTGTCTCTGACAG 5413  
DB |||||

QY 5590 CTCATCCGTGGTTGTATCAGGACTTATGGTGGCCACCACCAATATGAAGTGAGTGTCTATGC 5649  
DB |||||  
QY 5414 CTCATCCGTGGTTGTATCAGGACTTATGGTGGCCACCACCAATATGAAGTGAGTGTCTATGC 5473  
DB |||||  
QY 5650 TCTTAAGGACACTTTGA CAAGCAGACAGCTCAGGAGTTGTCAACA CTCTGAGGAATGT 5709  
DB |||||  
QY 5474 TCTTAAGGACACTTTGA CAAGCAGACAGCTCAGGAGTTGTCAACA CTCTGAGGAATGT 5533  
DB |||||  
QY 5710 CAGCCCAACAAGAGGGCTCTGTGACAGATCTACTGAGACCAACCATCACCATTAGCTG 5769  
DB |||||  
QY 5534 CAGCCCAACAAGAGGGCTCTGTGACAGATCTACTGAGACCAACCATCACCATTAGCTG 5593  
DB |||||  
QY 5770 GAGAAACAAGACTGAGACGATCACTGGCCCTCAAGTTGATGCCGTTCCAGCCAAATGGCCA 5829  
DB |||||  
QY 5594 GAGAAACAAGACTGAGACGATCACTGGCTTCCAAGTTGATGCCGTTCCAGCCAAATGGCCA 5653  
DB |||||  
QY 5830 GACTCCAATCCAGAGAACCATCAAGCCAGATGTCAAGAGCTACACCATCAGTCGGCTTACA 5889  
DB |||||  
QY 5654 GACTCCAATCCAGAGAACCATCAAGCCAGATGTCAAGAGCTACACCATCAGAGTTTACA 5713  
DB |||||  
QY 5890 ACCAGCACTGACTCAAGATCTA CCGTGTACA CCGTTGAATGACAATGCTCGGAGTCCCC 5949  
DB |||||  
QY 5714 ACCAGCACTGACTCAAGATCTA CCGTGTACA CCGTTGAATGACAATGCTCGGAGTCCCC 5773  
DB |||||  
QY 5950 TGTGGTCACTGACGCGCTCCACTGCCATTCGATGCACCATCCAACTGCGTTTCTCGGCCAC 6009  
DB |||||  
QY 5774 TGTGGTCACTGACGCGCTCCACTGCCATTCGATGCACCATCCAACTGCGTTTCTCGGCCAC 5833  
DB |||||  
QY 6010 CACACCAATTTCTTCTGGTATCATGCGCAGCGCCACCGTCCAGGATTAACCGGCTACAT 6069  
DB |||||  
QY 5834 CACACCAATTTCTTCTGGTATCATGCGCAGCGCCACCGTCCAGGATTAACCGGCTACAT 5893  
DB |||||  
QY 6070 CATCAAGTATGAGAGCCCTGGGTCTCTCCAGAGAGTGGTCCCTCGGCCCGGCGCTGG 6129  
DB |||||  
QY 5894 CATCAAGTATGAGAGCCCTGGGTCTCTCCAGAGAGTGGTCCCTCGGCCCGGCGCTGG 5953  
DB |||||  
QY 6130 TGTCAAGAGGCTTACTATTACTTGGCTCGAAACCGGAAACCGAATATACAAATTTATGTCA 6189  
DB |||||  
QY 5954 TGTCAAGAGGCTTACTATTCTTGGCTCGAAACCGGAAACCGAATATACAAATTTATGTCA 6013  
DB |||||  
QY 6190 TGGCTTGAAGAAATTAATCAGAGAGCGAGCCCTGTATGGAAGGAAAGAGAGAGAGCT 6249  
DB |||||  
QY 6014 TGGCTTGAAGAAATTAATCAGAGAGCGAGCCCTGTATGGAAGGAAAGAGAGAGAGCT 6073  
DB |||||  
QY 6250 TCCCAAACTGGTAAACCTTCCACACCCCAATTTCTATGGACAGAGATCTTGGATGTTCC 6309  
DB |||||  
QY 6074 TCCCAAACTGGTAAACCTTCCACACCCCAATTTCTATGGACAGAGATCTTGGATGTTCC 6133  
DB |||||  
QY 6310 TTCCACAGTTCAAAAGACCCCTTTTCTCACCACCCCTGGTATGACACTGGAAATGGTAT 6369  
DB |||||  
QY 6134 TTCCACAGTTCAAAAGACCCCTTTTCTCACCACCCCTGGTATGACACTGGAAATGGTAT 6193  
DB |||||  
QY 6370 TCAGCTTCTGGCACTTCTGGTCAGCAACCCAGTGTGGGCAACCAATGATCTTTGAGGA 6429  
DB |||||  
QY 6194 TCAGCTTCTGGCACTTCTGGTCAGCAACCCAGTGTGGGCAACCAATGATCTTTGAGGA 6253  
DB |||||  
QY 6430 ACATGGTTTTAGGCGGACCAACCCGCCACCAACCGGCCACCCCCATAAGGCAATAGGCCAAG 6489  
DB |||||  
QY 6254 ACATGGTTTTAGGCGGACCAACCCGCCACCAACCGGCCACCCCCATAAGGCAATAGGCCAAG 6313  
DB |||||  
QY 6490 ACCATACCCGCGAATGTAGGACAGAGCTCTCTCTCAGACCAACCATCTCATGGGCCCC 6549  
DB |||||  
QY 6314 ACCATACCCGCGAATGTAGGACAGAGCTCTCTCTCAGACCAACCATCTCATGGGCCCC 6373  
DB |||||  
QY 6550 ATTCCAGGACACTTCTGAGTACATCATTTTCTGTCATCTCTGTCAGTCTGATGAAGAAC 6609  
DB |||||  
QY 6374 ATTCCAGGACACTTCTGAGTACATCATTTTCTGTCATCTCTGTCAGTCTGATGAAGAAC 6433  
DB |||||  
QY 6610 CTTTACAGTTTCCCTGGAACTTTCTACAGTGCCACTCTGACAGGCTCACCAGAGG 6669  
DB |||||  
QY 6434 CTTTACAGTTTCCCTGGAACTTTCTACAGTGCCACTCTGACAGGCTCACCAGAGG 6493  
DB |||||  
QY 6670 TGCCACCTCAACATCATTAGTGGAGGCACTGAAAGACAGCAGAGGATTAAGGTTCCGGGA 6729  
DB |||||

6494 TGCACCTACACATCATAGTGGAGGCACTGAAGACACAGAGGCATAGGTTCCGGA 6553  
6730 AGAGGTTGTTACCGTGGGCAACTCTGTCAACGAAGCTTGAACCAACCTACGGATGACTC 6789  
6554 AGAGGTTGTTACCGTGGGCAACTCTGTCAACGAAGCTTGAACCAACCTACGGATGACTC 6613  
6790 GTGCTTTGACCCCTACACAGTGTCCCATTTATCCGTTGGAGATGAGTGGGAACGAATGTC 6849  
6614 GTGCTTTGACCCCTACACAGTGTCCCATTTATCCGTTGGAGATGAGTGGGAACGAATGTC 6673  
6850 TGAATCAGGCTTTAAACTGTGTGCCAGTGCTTAGGCTTTGGAAGTGGTCAATTCAGATG 6909  
6674 TGAATCAGGCTTTAAACTGTGTGCCAGTGCTTAGGCTTTGGAAGTGGTCAATTCAGATG 6733  
6910 TGAATCAGGCTTTAAACTGTGTGCCAGTGCTTAGGCTTTGGAAGTGGTCAATTCAGATG 6969  
6734 TGAATCAGGCTTTAAACTGTGTGCCAGTGCTTAGGCTTTGGAAGTGGTCAATTCAGATG 6793  
6970 CCGTCAGGAGAAAATGGCCAGATGATGAGTGTGCAATGCTTTGGAAACGGAAGAGAGA 7029  
6794 CCGTCAGGAGAAAATGGCCAGATGATGAGTGTGCAATGCTTTGGAAACGGAAGAGAGA 6853  
7030 ATTCAAGTGTACCCCTCATGAGGCAACGTTTACGATGATGGAAGACATACCACTAGG 7089  
6854 ATTCAAGTGTACCCCTCATGAGGCAACGTTTACGATGATGGAAGACATACCACTAGG 6913  
7090 AGAAGTGTGAGGAGAAAATGGCCAGATGATGAGTGTGCAATGCTTTGGAAACGGAAGAGAGA 7149  
6914 AGAAGTGTGAGGAGAAAATGGCCAGATGATGAGTGTGCAATGCTTTGGAAACGGAAGAGAGA 6973  
7150 GCGGGGCTGCGCTGTGACAACTGCGCGAGACCTGGGGGTGAACCCAGTCCCGAAGGCAC 7209  
6974 GCGGGGCTGCGCTGTGACAACTGCGCGAGACCTGGGGGTGAACCCAGTCCCGAAGGCAC 7033  
7210 TACTGCGCAGTCTCAACACAGTATTTCTCAGATACCACTCAGAGAACAACTAATGT 7269  
7034 TACTGCGCAGTCTCAACACAGTATTTCTCAGATACCACTCAGAGAACAACTAATGT 7093  
7270 TAATGCCCAATAGTGTTCATGCTTTAGATGATGAGGCTGACAGAGATTCCTCG 7329  
7094 TAATGCCCAATAGTGTTCATGCTTTAGATGATGAGGCTGACAGAGATTCCTCG 7153  
7330 AGAGTAACTCATCTTCCATCCAGAGAACAGCATGCTCTCTCCCAAGATCCATGTA 7389  
7154 AGAGTAACTCATCTTCCATCCAGAGAACAGCATGCTCTCTCCCAAGATCCATGTA 7213  
7390 AACTGGAGTGAATGATGACAGACCCAGCTTACGATGCTTTCTTTTAAAGCCCTTTGC 7449  
7214 AACTGGAGTGAATGATGACAGACCCAGCTTACGATGCTTTCTTTTAAAGCCCTTTGC 7273  
7450 TCTGGGAGAGTTCTCCAGCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 7509  
7274 TCTGGGAGAGTTCTCCAGCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 7333  
7510 GTTTCCTGAGGCTTTCTCATAATGAGGCTGACATTCGCTTTGCTGCTTGAAGTAT 7569  
7334 GTTTCCTGAGGCTTTCTCATAATGAGGCTGACATTCGCTTTGCTGCTTGAAGTAT 7393  
7570 TCAATACCGCTCAGATTTTAAATGAGTGAATTTGATTTGATTTGATTTGATTTGATTTGAT 7629  
7394 TCAATACCGCTCAGATTTTAAATGAGTGAATTTGATTTGATTTGATTTGATTTGATTTGAT 7453  
7630 AAGCATATGACGCAACCAAGATGCAATTTTGAATGATATGACCAAAATTTTAACT 7689  
7454 AAGCATATGACGCAACCAAGATGCAATTTTGAATGATATGACCAAAATTTTAACT 7513  
7690 AGGAAGTCAACCAACCACTTCTTCACTTAACTGATGCTGCGCAATCTGATAGGA 7749  
7514 AGGAAGTCAACCAACCACTTCTTCACTTAACTGATGCTGCGCAATCTGATAGGA 7573  
7750 ACAAGCATGATCTTGTACTGTGATATTTTAAATATCCACAGTACT 7795

Db 7574 ACAAGCATGATCTTGTACTGTGATATTTTAAATATCCACAGTACT 7619  
RESULT 14  
US-09-917-800A-1731  
; Sequence 1731, Application US/09917800A  
; Patent No. US20020119462A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Castle, Arthur  
; APPLICANT: Blashoff, Michael  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5038-US  
; CURRENT APPLICATION NUMBER: US/09/917,800A  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,040  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,880  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: US 60/290,029  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US 60/290,645  
; PRIOR FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US 60/292,336  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/295,798  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/297,457  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,884  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,459  
; PRIOR FILING DATE: 2001-07-09  
; NUMBER OF SEQ ID NOS: 1740  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1731  
; LENGTH: 8329  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_019143  
US-09-917-800A-1731

Query Match 68.3%; Score 5320.6; DB 9; Length 8329;  
Best Local Similarity 81.8%; Pred. No. 0; Indels 397; Gaps 14;  
Matches 6484; Conservative 0; Mismatches 1044;  
QY 184 GGAATTTTCCCGCAGGTTGCGAAGGGAAGCAAACTTGGTGGCAACTTGCCTCCCGGTGCG 243  
Db 123 GGACTTTTCTGTCAGGCTGCGAGGGAA-CCAACTTCTGGCCACTAGCTCTCTGGAGAG 181  
QY 244 GGGTCTCTCC-CCACCGTCTCAACATGCTTAGGGTTCGGGGCCG---GGCTGCTGC 299  
Db 182 GGCAGCTCTCTCCACTCCACTCAAGATGCTCAGGGTTCGGGACCCGGCGGCTGCTGC 241  
QY 300 TGCTGGCGCTCTCTGCTGGGACAGCGTGCCTCCACGGGAGCCTCGAAGAGCAAGA 359  
Db 242 TGCTAGCAGTCTCTGCTGGGACATCGGTGCGCTGCACCGAAACCGGAAGAGCAAGA 301  
QY 360 GGCAGGCTCAGCAAAATGGTTCCAGCCCGAGTCCCGGTGGCTGTGCTCAGTCAAGAGCCCG 419  
Db 302 GGCAGGCTCAGCAAAATCGTCAGCCTCCGTCGGGCTGCTGCTCAGTCAAGAGCCTG 361  
QY 420 GTTGTATGACAAATGGAACACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 479  
Db 362 GCTGTTTGAACAGGAGCAATTAATCAGATAATCAGCAGTGGGACGACCTACCTAG 421  
QY 480 GCAATGCGTTGCTTGTACTTGTATGGAGAACCCAGAGTTTAACTGCGAAAGTAAAC 539  
Db 422 GCAACGCCCTGGTTTGTACCTGCTATGGAGAACAGAGGTTTAACTGCGAGAGCAAGC 481

540 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 599  
 Db |  
 482 CTGAACCTGAAGAGACCTGTTTTCACAAATACACTGGAAACACTTACAAAGTGGGTGACA 541  
 600 CTTATGAGCGCTTAAAGACTCCATGATCTGGGACTGTACTCTGCATCGGGCTGGGGCGAG 659  
 Db |  
 542 CTTATGAGCGCTTAAAGACTCCATGATCTGGGACTGTACTCTGCATGGGGCTGGGGCGAG 601  
 660 GGAGATAAGCTGTACCATCCAAACCGCTGCCATGAAGGGGGTCACTCTCAAAAGATTG 719  
 2y |  
 602 GCAGATCAGCTGTACCATTCGCTGCCATGAAGGGGGTCACTCTCAAAAGATTG 661  
 Db |  
 720 GTGACCTGGAGAGACCAATGAGACTGGTGTATCATGTTAGAGTGTGTGTTCTTG 779  
 2y |  
 662 GTGACAGTGGAGAGGCCACATGAGACTGGTGTATGTTGGAGTGTGTGTTCTTG 721  
 Db |  
 780 GTAATGGAAAAGGAGATGGAACCTGCAAGCCCATAGCTGGAAGTGTGTTGATCATGCTG 839  
 2y |  
 722 GGAATGGAAAAGGGAATGGACCTGCAAGCCCATAGCTGGAAGTGTGTTGATCACGCTG 781  
 Db |  
 840 CTGGGACTTCTATGTGTCGAGAAACGTCGGGAGAGCCCTACCAAGGCTGGATGATGG 899  
 2y |  
 782 CTGGGACTTCTATGCTGCTGGGGAGAGCCTGGGAAAGCCCTACCAAGGCTGGATGATGG 841  
 Db |  
 900 TAGATTGTACTTGGCTGGGAGAGGAGCGAGCGCATCTTGCACTTCTAGAAATAGAT 959  
 2y |  
 842 TGGACTGTACTTGTCTGGGCGAAGGCAATGGGCGTATCACCTGCACCTCCCGGAACAGAT 901  
 Db |  
 960 GCAACGATCAGGACAAAGGACATCTCTATAGAAATGGAGACACCTGGAGCAAGAGGATA 1019  
 2y |  
 902 GCAATGATCAGGACAAAGGACGCTCTACAGAAATGGAGACACATGGAGCAAGAGGACA 961  
 Db |  
 1020 ATCAGAGGAAACCTGCTCAGTGTGCACTGTCACAGGCAACCGGCCGAGGAGAGTGGAAAGTGTG 1079  
 2y |  
 962 ACAGAGGAAACCTGCTCAGTGTGCTGTCACAGGCAACCGGCCGAGGAGTGGAAAGTGTG 1021  
 Db |  
 1080 AGAGGCAACCTCTGTGAGAGACCATGACGAGGAGATCTGGGCCCTTACCGATGTTGCTG 1139  
 2y |  
 1022 AGCGACA---TGTTCTACAGAGTGTCTCAGCTGTGATCTGGCTCTCTCACAGATGTCGAA 1078  
 Db |  
 1140 CAGCTGTTTACCAACCGCAGCTCACCCCGCAGCTCTCTCCCTATGCGCACCTGTGTACAG 1199  
 2y |  
 1079 CAGCTATTTTACCAACCGCAGACCCCGCAGCTCTCTCCCTATGCGCACCTGTGTACAG 1138  
 Db |  
 1200 ACAGTGTGTGTCTCTGTGGGATGCAAGTGTGAGGACCAAGGAAATAGCAAA 1259  
 2y |  
 1139 ACAGCGTGTGTCTCTGTGGAAATGCAAGTGTGAGTCTCAAGGAGACAAGCAGA 1198  
 Db |  
 1260 TGCTTTGACGCTGCTGGGCAACGAGTCACTGTCGCAAGAGACAGCTGTACCCAGACTT 1319  
 2y |  
 1199 TGCTGTGCACTTGGCTGGGCAATGGCGTCACTGTCGCAAGAGACAGCTGTACCCAGACTT 1258  
 Db |  
 1320 ACGGTGGCACTCAAAATGGAGAGCAATGTCTTACCACTTCACTACCAATGGCAGGAGCT 1379  
 2y |  
 1259 ACGGTGGCACTCAAAATGGGAGAGCTGTGTCTCTCCCTTCACTACCAACGCTAGGAGCT 1318  
 Db |  
 1380 TCTACTCTGCAACAGAGAGGCGACAGGAGGAGTCTTGGTGGAGCACTTCTGCAAGTTCGA 1439  
 2y |  
 1319 TCTACTCTGCAACAGGAGGCGGAGGAGGAGTCTGTGGTGTAGCACTTCTCAA 1378  
 Db |  
 1440 ATTATGAGCAGGACCAAGAAATACTCTTTCTGCAAGAGACCACTGTTTGTGTTGAGACTC 1499  
 2y |  
 1379 ATTATGACACAGACAGAGATTTCTTCTGCAAGACCAACGCGGCTTTTGTGTTGAGACTC 1438  
 Db |  
 1500 GAGGAGGAAATTCGAATGGTGGCTTGTGGCACTTCTCCCTTCTATACAAACACCAAT 1559  
 2y |  
 1439 GAGGTGGAAATTCGAATGGTGGCTTGTGGCACTTCTCCCTTCTGACAGCAACCGGAAT 1498  
 Db |  
 1560 ACACCTGATTGCACTTCTGAGGCGAGAGAGCAACATGAAGTGGTGGGAGCAGACAGA 1619  
 2y |  
 1499 ACAGGCACTGTACTTCTGAGGGTAGGGGAGCAACATGAAATGGTGGGCAACCCAGAA 1558  
 Db |

1620 ACTATGTCGCGACAGAGTTTGGGTTCTGCCCATGGCTGCCACAGGAAATCTGCA 1679  
 2y |  
 1559 ACTAGATGCCCATCAGAAAGTTTGGATTCTGCCCATGGCTGCCCATGAGAGATCTGCA 1618  
 Db |  
 1680 CAAACAAATGAAGGGTCAATGTACCCCATTTGAGATTCAGTGGGATAAGCAGCATGACATGG 1739  
 2y |  
 1619 CGACCAACGAAGGGTCAATGTATCGCATTTGGGACCACTGGGATAAGCAGCATGACCTGG 1678  
 Db |  
 1740 CTCACATGATGAGGTGCACGTGTGTTGGGAATGTCGTGGGGAATGGACATGCAATGGCT 1799  
 2y |  
 1679 GCCACATGATGAGGTGCACGTGTGTTGGGAACGGCCGTGGACAATGGGCCCTTGCAATCCCCT 1738  
 Db |  
 1800 ACTCGCAGCTTCGAGATCAGTGCATTTGTTGATGACATCACTTACAAATGTGAACGACAT 1859  
 2y |  
 1739 ACTCCAGCTTCGAGATCAGTGCATTCGTTGATGACATTTACTTACAACTGCAACGACAT 1798  
 Db |  
 1860 TCACAGAGCTCATGAAGGGGACATGCTGAATCTGATGCTTCCGTTCAGGTCAGGTCGGG 1919  
 2y |  
 1799 TCACAGAGCTTCAGAGGAGGACATGCTGAATCTGCTTCCGTTCAGGTCAGGTCGGG 1858  
 Db |  
 1920 GCAGGTGGAAGTGTGATCCCGTCGACCAATGCGAGATTTCAGAGACTGGGAGCTTTTATC 1979  
 2y |  
 1859 GCAGATGGAATGTGACCCCATCGACCGATGCCAAGATTCAGAGACCCGACATTTTACC 1918  
 Db |  
 1980 AAATTGGAATTCATGGGAGAGATGTCATGCTGTCAGATCAGATCAGTCTACTGCTATG 2039  
 2y |  
 1919 AGATTGCTGACTCCTGGGAGAAAGTTTGTGTCATGCTGTCAGATCAGATGTTACTGTTAG 1978  
 Db |  
 2040 GCGTGGCAATTCGGGAGTGGCAATTCGCAACCTTTTACAGACCTTATCAAGCTCAAGTGTG 2099  
 2y |  
 1979 GCGTGGCAATTCGGGAGTGGCAATTCGCAACCTTATCAAGCTTATCAAGCTCAAGTGTG 2038  
 Db |  
 2100 CTGTGCAAGTATTTATCACTGAGACTCGAGTTCAGGCAACTCCCAACCCCATCCAGTGA 2159  
 2y |  
 2039 CTGTTCAGTAAATATATACGAGAGACCCCGCAGGAGGAGTTCACCAACCCCATCCAGTGA 2098  
 Db |  
 2160 ATGCACCAACAGCCTCTCACTTTCCAGTATCACTTCAGGTGGAGACTTAAATTTCTG 2219  
 2y |  
 2099 ATGCCCGGAGCTTCAACATCAACAGTACATTTCTAGGTGGAGACTTAAATTTCTG 2158  
 Db |  
 2220 TAGGCGGTTGGAAGAGCTTACATACCAAGGCACTTAACTCTTACACCATCAAGGCT 2279  
 2y |  
 2159 CGGCTCGCTGGAAGAGCTACATTCAGGAGGCACTTAACTCTTATACCATCAAGGCT 2218  
 Db |  
 2280 TGAAGCTGCTGTGTATACGAGGCGCAGCTCATCAGCATCCAGAGTACGCGCCACCAAG 2339  
 2y |  
 2219 TGACCCCGAGTGTGTATACGAGGAGAGCTCATCAGCATCCAGAGTACGCGCCACCAAG 2278  
 Db |  
 2340 AAGTGACTCGCTTGAACCTTACCAACCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2399  
 2y |  
 2279 AAGTGACTCGCTTGAACCTTACCAACCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2335  
 Db |  
 2400 TGAAGAGAGAGGAGTCCCTTTTCTCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2459  
 2y |  
 2336 TGAAGAGAGAGTCCCTTTTCTCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2395  
 Db |  
 2460 TCACAGCCTGAGTGTGTGTCCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2519  
 2y |  
 2396 TCACAGCCTGAGGAGTGTGTGTCCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2455  
 Db |  
 2520 GGTGGAATATGAGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2579  
 2y |  
 2456 GAGTGGAGTACCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2515  
 Db |  
 2580 CAGGCACTTCTGTGAACATCCCTGACCTCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2639  
 2y |  
 2516 CAGGCACTTCTGTGAACATCCCTGACCTCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2575  
 Db |  
 2640 ATCAGATATCTGAGGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2699  
 2y |  
 2576 ATCAGATATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2635  
 Db |  
 2700 CTGATGCCCTCTCTGAGCAGGAGTGTGGACCAAGTTGATGACACCTCAATTTGTTGCTGCT 2759  
 2y |



b	2636	CTGATGGCGCTCCAGACCCCTACTGTGTGGACAGAGTTGATGACACTTCCTCATTTGTTTCGAT	2695
y	2760	GGACGACAGCCCGAGCTCCCATCACAGGGTACAGAAATAGTCTATTTCGCCATCAGTAGAAG	2819
b	2696	GGACGACAGCCCGAGCACTTATCACAGGGTACAGGATTTGCTATTTCACCTTCAGTAGAAG	2755
y	2820	GTAGCAGCAGAACTCAACTTCCTGTAAGCTGCAAACTCCGTCACGCTCAGTGACTGC	2879
b	2756	GCAGTAGCAGAACTCAACTTCCTGTAAGCTGCAAACTCCGTCACGCTCAGTGACTGC	2815
y	2880	AACCTGCTGTTTCAGTATAACATCACTATCTATGCTGTGGAAGAAATCAAGAAAGTACAC	2939
b	2816	AGCCCGGTGTTTCAGTACAACATCACTATCTATGCTGTGGAAGAAATCAAGAAAGTACAC	2875
y	2940	CTGTTGTCATTCAACAGAAACCACTGGGACCCGCTCAGATACAGTGCCCTCCCA	2999
b	2876	CCGTTTTCATCCAGCAGGAGACTACTGGGCTCCACGATCCGATGATGTTCCCGCTCCAA	2935
y	3000	GGGACCTGCGATTTGTGGAAAGTGACAGCGTGAAGTCAACCATCATGTGGACACCCGCTG	3059
b	2936	AGGACCTACAGTTTGTGGAAAGTGACAGCGTGAAGTCAACCATCATGTGGACACCTCCTA	2995
y	3060	AGAGTGCAGTACCGGCTACCGTGTGATGTATCCCGGTCAACCTGCCTGGCGGACAG	3119
b	2996	ATAGCGGAGTACCTGGATACCGTGTGATGTCTGCTGTCAACCTGCCAGGGGAACATG	3055
y	3120	GGCAGAGCTGCCATCAGCAGGAAACCTTTTGAGAAAGTCAACCGGCTGTCCCTGGGG	3179
ib	3056	GGCAGAGCTGCCCTGTCAACAGGAAACCTTTGCTGAAGTCAACCGGCTGTCCCGAGGGG	3115
iy	3180	TCACCTATTACTTCAAGTCTTTTCAGTGTGATGTATCCCGGTCAACCTGCCTGGCGGACAG	3239
ib	3116	TCAGTGACTCTTTCAAGTCTTTGCTGTGATCATAGGCGAGGAAAGCAAGCTCTACAG	3175
iy	3240	CTCAACGACAAACCAACTGGATGCTCCCACTAACCTCCAGTTTGTCAATGAACACTGATT	3299
ib	3176	CACAACGACCAACCAACTCGATGCTCCCACTAACCTCCAGTTTGTCAATGAACAGGACA	3235
iy	3300	CTACTGCTCGTGAGTGAAGTCACTCCGCTGGGCCCCAGATAACAGGATACCGACTGACCG	3359
ib	3236	GAACAGTTCTGGTAACTTGGACTCACTCCGAGCCCGGATAGCAAGGTACCGACTGACAG	3295
iy	3360	TGGGCTTACCCGAGGAGCAGCCGAGGAGTCAATGTGGGTCCTCTGTCTCCAAAGT	3419
ib	3296	TGGGCTTACCCGAGGAGGAGCAGCCGAGGAGTCAATGTGGGTCCTCTGTCTCCAAAGT	3355
iy	3420	ACCCACTCAGGAGTCTGAGGCTGATCTGAGTACACCGTATCCCTCGTGGCCATAAAGG	3479
ib	3356	ATCCCTCTGAGAAATCTGACGCTGGTCTGAGTACATGTGACCTTGATGGCTGTGAAAG	3415
iy	3480	GCAACCAAGAGAGCCCAAGCCACTGGAGTCTTTAACAACATGCAAGCTGGGAGTCTTA	3539
ib	3416	GCAACCAAGAGAGTCCCAAGCCACTGGAGTCTTTAACAACATGCAAGCTGGGAGTCTTA	3475
iy	3540	TTCCACCTTACAAACCCAGGTGATGAGACCACTTTGATCATCATGAGCGCTGCTC	3599
ib	3476	TTCCACCTTATAACCAAGAGGTGATGAGACCACTTTGATCATCATGAGCGCTGCTC	3535
iy	3600	CAAGAAATGGTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGAGCCACGAGAG	3659
ib	3536	CAAGAAATGGTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGTGAAGCAGCCGAGAG	3595
iy	3660	TGACCTTCAGACTCAGGAGGAGTCTGTTGTGCTGGGCTTGACTCCAGGAGTAGAATACGCT	3719
ib	3596	TGACCTTCAGACTCAGGAGGAGTCTGTTGTGCTGGGCTTGACTCCAGGAGTGAATACAGT	3655
iy	3720	ACACCATCAAGTCTGAGATGAGACAGGAAAGAGATGGCCCAATTTGTAAACAAAGTGG	3779
ib	3656	ACACCATCAAGTCTGAGGAGGAGGAGGAGGAGATGCAACCAATTTGTAAACAAAGTGG	3715
iy	3780	TGACACCATTTGTTCCCAACCAACTTTGATCTTGAGGAGCAAAACCTTGACACTGGAGTGC	3839



2y	4647	CCCCACAGCCTACTGATCAGCTGGGATGCTCTCTGTGTACAGTGAGATATTTACAGGA	4706	5727	CTCTGTGTGACAGATGCTACTGAGACCAACCATCAACATTTAGCTGGAGAACCAAGACTGAGA	5786
Dd	4856	CCCCACAGCCTGCTCATAGTTGGAAACCCCGCGCTCTCTGTGGCTATTACAGAA	4915	5936	CCCGTGTGACCGACGCTACAGAACTACCATCACTATTAGCTGGAGAACGAGACAGAGA	5995
2y	4707	TCACCTTACGAGAAACAGGAGGAATAGCCTGTCCAGGAGTTCACTGTGCCCTGGAGCA	4766	5787	CGATCACTGGCTCCAAAGTTGATGCGGTTCCAGGCAATGGCCAGACTCCAATTCAGAGAA	5846
Dd	4916	TCACCTATGGAGACAGGAGGAATAGCCTGTCCAGGAAATTCACCTGTCCCGGAGCA	4975	5996	CGATCACTGGCTTCCAAAGTTGATGCGGTTCCAGGCAATGGCCAGACCCCGGTTACAGGA	6055
2y	4767	AGCTACAGCTACCATCAGCGGCTTAAACCTGGAGTTGATTATACCATCACTGTGTATG	4826	5847	CCATCAAGCCAGATGTGAGAAGCTACACCATCACTGGGTTTAAACCCAGGCACTGACTACA	5906
Dd	4976	AGTCCACCGCCACCATTAACAAATTTAAACGAGAGCAGACTACACCATCACTGTGTATG	5035	6056	CCATCAGCCCGGATGTGAGAAGCTATATACTATTACAGGTTTACAGCCAGGCACTGACTACA	6115
2y	4827	CTGTCACTGGCGGTGGAGACAGCCCGCAAGACAGCAACCAATTCATTAAATTAACGAA	4886	5907	AGATCTACTGTGTACACTTGAATGACATGCTCCGAGCTCCCTGTGTGTATCATCAGCGCT	5966
Dd	5036	CTGTCACTGGCGGTGGAGACAGCTCCAGCAGCAGCAAGCCAGTTTCCATCAATTAATCAAA	5095	6116	AGATCCACTGTGTACAGCTCAACGAAATGCCGAGCTCTCTGTGTGTATTGATGCT	6175
2y	4887	CAGAAATGACAAACCATCCAGATGCAAGTACCGGATGTTTCAGGACAAACAGATTAGTG	4946	5967	CCATGCGCAATGATGACACCATCCAACTGCGTTTCTGGCCACACACACCCAAATTCCTTGC	6026
Dd	5096	CAGAAATGACAAACCATCCAGATGCAAGTACCGGATGTTTCAGGACAAACAGATTAGTG	5155	6176	CCACGGCCATTGATGCCCATCCAACTGCGTTTCTGGCCACACACACCCAACTCCTTGC	6235
2y	4947	TCAAGTGGCTGCTTCAAGTTCCCTGTACTGTGTTACAGAGTAAACCACTCCCAAAA	5006	6027	TGGTATCATGGCAGCGGCACGTCGAGGATTACCGGTACATCATCAAGTATGAGAGC	6086
Dd	5156	TCAGTGGCTGCTTCAAGTTCCCTGTACTGTGTTACAGAGTAAACCACTCCCAAAA	5215	6236	TGGTATCATGGCAGCGGCACCGGTGCGGATTTACTGGCTTACATTATCAAGTATGAGAGC	6295
2y	5007	ATGACACGAGCAACAAATTAACCTGAGTCCAGTCCAGATCAACAGAAATGACTATTG	5066	6087	CTGGGTCTCTCCACAGAGAAAGTGGTCCCTCGGCCCGCCCTGGTGTACAGAGGCTACTA	6146
Dd	5216	ATGGCTTAGGACCAACAAATTAACCTGAGTCCAGTCCAGATCAACAGAAATGACTATTG	5275	6296	CTGGATCCCTCCACAGAGAAAGTGGTCCCTCGGCCCGCCCTGGTGTACAGAGGCTACTA	6355
2y	5067	AAGGCTTCAGCCACAGTGGAGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5126	6147	TTACTGGCTCGAAACCGGAAACCGAAATATACATTTATGTATTTGCTTCCCTCGGATATATC	6206
Dd	5276	AAGGCTTCAGCCACAGTGGAGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5335	6356	TCATGTGTGTGAGCAGCAACCGAGTACACCATCTATGTATGTCATGCACTGAGAAACAATC	6415
2y	5127	AGAGTACGCTCTGTTTACAGTACGAGTAAACCAATGATCGCCCTTAAGGACTGGCAT	5186	6207	AGAAGAGCAGGCGCTGTGATTGGAAGGAAAGACAGAGAGCTTCCCAACTGGTAAACC	6266
Dd	5336	AAAGCAGGCTGCTGTTTACAGTACGAGTAAACCAATGATCGCCCTTAAGGACTGGCAT	5395	6416	AGAAGAGTGGGCGCTGATTGGAGGAAAGACAGAGTGGCTTCCCAACTGGTAAACC	6475
2y	5187	TCAGT	5246	6267	TTCCACACCCCAATCTTATGAGCAGAGATCTTGGATGTTCCTTCCAGTGTCAAAGA	6326
Dd	5396	TCAGT	5455	6476	TTCCACACCCCAATCTTATGAGCAGAGATCTTGGATGTTCCTTCCAGTGTCAAAGA	6535
2y	5247	CCAGGTACAGGCTGACCTACTCGAGCCTGAGGATGGAATCCATGAGTATTCCTTGCAC	5306	6327	CCCTTTCTGTGACCCACCTGGGTATGACACTGGAATGGTATTTCAGTGTCTCTGCACTT	6386
Dd	5456	CCAGGTACAGGCTGACCTACTCGAGCCTGAGGATGGAATCCATGAGTATTCCTTGCAC	5515	6536	CCCTTTCTGTGACCCACCTGGGTATGACCGGAAATGGTATTTCAGTGTCTCTGCACTT	6595
2y	5307	CTGATGTTGAAGAGACACTCGAGGCTGCAAGCCTCAGACCGGTTCTGAGTACACAG	5366	6387	CTGGTCAGCAACCCAGTGTGGGCAACAAATGATCTTTAGAGAAATGTTTGGGCGGA	6446
Dd	5516	CTGATGTTGAAGAGACACTCGAGGCTGCAAGCCTCAGACCGGTTCTGAGTACACAG	5575	6596	CCACCAACAAACCCAGTGTGGGCAACAAATGATCTTTAGAGAAATGTTTGGGCGGA	6655
2y	5367	TCAGT	5426	6447	CCACACCGCCACAGCGGCCACCCCATTAAGGATAGGCAAGACCATACCCGCGGAATG	6506
Dd	5576	TCAGT	5635	6656	CCACGCCACCCACTGGGCCACCCCGCTGAGGCTTAGGCCAAGACCATACCTGCCGAATG	6715
2y	5427	CAGCTATTCTGCAACCACTGACCTGAGTTTCACTCAGGTACACCCACAGCCTGAGCG	5486	6507	T-----T-----T-----T-----T-----T-----T-----T-----	6507
Dd	5636	CAGCAATTCCTGCGCAACCAATCTGAAAGTTCACTCAGGTGTACCCACCTTGTACTG	5695	6716	TAGATGAGGAGTGCATAATCGGTGATGTTCCAGGGAGACGTAGACTACCACTCTATC	6775
2y	5487	CCAGTGGACACCAACCAATGTTTCAAGTCTCAGTGTGATATGAGTGGGTTGACCCCAAG	5546	6508	-----AGGCAAGAGCTCTCTCTCAGAGAA	6533
Dd	5696	CCAGTGGACACCAACCAATGTTTCAAGTCTCAGTGTGATATGAGTGGGTTGACCCCAAG	5755	6776	CTCATGTTCCGGGTCAATCCAAATGCCCTTACAGGACAAAGAGCTCTCTCTCAGAGAA	6835
2y	5547	AGAAGACCGGACCAATGAAGAAATCAACCTTGTCTGAGCCTCATCGTGGTGTGTAT	5606	6534	CCATCTCATGGGCCCCCATTTCCAGGACACTTTTGGAGTACATCAATTTTCATGTCTCTGTG	6593
Dd	5756	AGAAGACCGGACCAATGAAGAAATCAACCTTGTCTGAGCCTCATCGTGGTGTGTAT	5815	6836	CCATCTCTTGGACGCCATTTCCAGGAGAGTTCTGAGTACATCATTTTCATGCCAACCTGTG	6895
2y	5607	CAGGACTTATGTTGCCACCAATATGAAGTGTGTCTATGCTCTTAAGGACACTTTGA	5666	6594	GCACTGATGAAGAACCTTTACAGTTTCCAGGTTCTGGAACTTCTACAGTGGCACTCTGA	6653
Dd	5816	CAGGCTCATGTTGGCCACCAATGATGAAGTGTGTCTATGCTCTCAAGGACACTTTGA	5875	6896	GCACTGACGAAGAGCCCTTACAGTTTCCAAAGTTTCCGGAAGTCTTCTACAGTGGCACTGA	6955
2y	5667	CAAGCAGACCACTCAGGAGTTGTACCACTCTGGAGAAATGTACGCCCAACAGAGAGG	5726	6654	CAGGCTCTACAGAGTGCACCTACCAATCATAGTGGAGGCACTGTAAAGACCAAGCA	6713
Dd	5876	CAAGCAGACCACTCAGGAGTTGTACCACTCTGGAGAAATGTACGCCCAACAGAGAGG	5935	6956	CTGGCTTACAGAGGGGTCACTTACCAATCATAGTGGAGGCTCTGACCAACAGAGGA	7015
2y	5876	CAAGCAGACCACTCAGGAGTTGTACCACTCTGGAGAAATGTACGCCCTCCAAAGAGG	5935	6714	GGCATAAGGTTCCGGGAAGAGGTTGTTTACCGTGGGCAACTCTCTCAACGAAGGCTTGAACC	6773

```

b 7016 GGCACAGGTCGCGAAGAGGTTGTTACTGTAGGCAACACTGTCAACGAAGCCCTGAACC 7075
y 6774 AACCTAGGATGACTGCTGCTTGACCCCTACACAGTGTCCCATTTATGCCGTTGAGATG 6833
b 7076 AGCCTAGGATGACTCATGCTTTGACCCCTTACACGGTTTCCCATACGCCGTTGAGAGG 7135
y 6834 AGTGGGAAAGAAATGCTGAATCAGGCTTTAAACTGTTGTGCCAGTGCCTTAGCTTTGGAA 6893
b 7136 AATGGGAGCGGTATCTGACTCTGGCTTTAAAGCTCACTTGGCAGTGTGGCTTTGGCA 7195
y 6894 GTGGTCATTTAGATGTGATTCATCTAGATGTTGTCATGACACAAATGGTGTGAACACTAAGA 6953
b 7196 GTGGTCATTTAGATGCGATTCATCTAAATGGTGGCATGACAAACGGTGTCAACTAAGA 7255
y 6954 TTGCAGAGAAGTGGGACCGTCAGGAGAAATATGGCCAGATGATGAGCTGCACATGCTTTG 7013
b 7256 TCGGAGAGAGTGGGATCGTCAGGAGAAATATGGCCAGCGGATGAGCTGCACATGCTCG 7315
y 7014 GGAACGGAAAGAGAAATTCAGTGTGAACCTCATGAGGCAACGTTTACGATGATGGA 7073
b 7316 GGAATGGAAGGAGAAATTCAAATGCGATCCCCATGAAGCAACGTTTATGACGACGGA 7375
y 7074 AGACATACACGCTAGGAGAACAGTGCAGAGGAATATCTCGTGTGCAATTTGCTCTGCA 7133
b 7376 AGACCTACACGCTAGGAGAACAGTGCAGAGGAATATCTCGAGGCAATTTGCTCTGCA 7435
y 7134 CATGCTTTGAGGCCAGCGGGCTGGCGTGTGACAACTGGCCGAGACCTGGG---GGTG 7190
b 7436 CGTGTTCGGGGGCCAGCGGGCTGGCGTGTGACAACTGGCCGAGACCTGGGGCTGTG 7495
y 7191 AACCCAGTCCGAGGACACTACTGCGGAGTCTTACAAACAGTATCTCAGAGATACCATC 7250
b 7496 AACCCAGTCCGAGTGTACCACTGGCCACACCTACAAACAGTATACAGAGATACCATC 7555
y 7251 AGAGAACAAACACTAATGTTAAATGGCCAAATGAGTGTCTTATGATGATGATGATG 7310
b 7556 AGAGAACAAACACTAATGTTAAATGGCCAAATGAGTGTCTTATGATGATGATGATG 7615
y 7311 CTGACAGAGAGATCCGAGAGTAATCATCTTTCCAAATCCAGAGGAACAAGCATGCT 7370
b 7616 CTGACAGAGATGATCCAGAGATGATCTTTCCATCCAGCCCAAGCAACAG--TGTC 7673
y 7371 CTCTGCCAGATCCATCTAACTGAGTGTGATGTAGCAGAC---CCAGCTTAGAGTCTT 7427
b 7674 CTCTACCAAGGTCAATCCACACCCAGTGTGATGTAGCAGACCTCCATTTCTGAGTGT 7733
y 7428 CTTTCTTTCTTAGCCCTTTGCTCTGAGGAGTCTCTCCAGTTCAGCTCAACTCACAGC 7487
b 7734 ATTTCAACCTTAAGCCCTTCTGCTGAGTCAAGTCTCAGCTTTCAGCTTTCAGCTT 7793
y 7488 TTCTCCAAGCATCACCTG--GGAGTTTCTGAGGTTTTTCTCAATAATGAGGCTGCACA 7546
b 7794 TTCTCCAAGCATCGCCCGCGGATGTTTGGAGCTTCCCTCTTAAATGGTGACATGG 7853
y 7547 TTGCTGTTCTGCTTGAAGTATTCATACCGCTCAGTATTTTAAATGAGTATCTTAA 7606
b 7854 TGCCCTGTTCTGCTTCAAGGATTTCAAGTCTGCTCAGTATTTATGTTCAAGAGATCAA 7913
y 7607 -----GATTTGTTTGGGATCAATAGGAAGCATATGACCAACCAAGATGCAAA 7657
b 7914 AGTTCTTGTGATTTGGTCTGGATCAA--AGGAAACACAGGTAGTACCAACCAAGATG 7971
y 7658 TGTTTGAATATGATGACCAAAATTTTAAATAGTAGGAAGTACCCAAACACTTCTGCTTT 7717
b 7972 TGAATGAATGTAGTACCAAGAGCGGAGCGGAAGTTTAAACCAAGACAGTCTCTGCTTT 8031
y 7718 CACTT 7722
b 8032 CTTT 8036

```

RESULT 15

```

US-10-191-803-113
; Sequence 113, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 8329
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_019143
US-10-191-803-113

```

```

Query Match      68.3%; Score 5320.6; DB 15; Length 8329;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 6484; Conservative 0; Mismatches 1044; Indels 397; Gaps 14;

QY 184 GGATTTTCCGCGAGGTTGGAAGGAGCAAACTTGGTGGCACTTCCTCCCGGTGCG 243
DB 123 GGACTTTTCTGTCAGGCTGCGAGGGAA--CCAACTTGTGGCCACTAGCTCTCTGGAGAG 181
QY 244 GGCTGCTCTCC--CCACCGCTCTCAACATGCTTAGGGGTCCGGGGCCCG---GGCTGCTGC 299
DB 182 GGCACTCTCTCTCCATCCACTCAAGATGCTCAGGGGTCCGGGACCCCGCGCGCTGCTGC 241
QY 300 TGCTGGCGCTGCTGTGCTGGGACAGCGGTGCCCTCCAGGGAGCTTCGAAGAGCAAGA 359
DB 242 TGCTAGCAGTCTGTGCTGGGACATCGGTGGCTGCACCCGAAACCGGGAAGAGCAAGA 301
QY 360 GGCAGGCTCAAGAAATGTTTCAGCCCGAGTCCCGGTGGGTGTGTCAGTCAAAAGCAAGCCCG 419
DB 302 GGCAGGCTCAAGAAATGTTTCAGCCCGTCCGTCGGCTGCTGTCAGTCAAGCAAGCCCTG 361
QY 420 GTTGTATGACAAATGGAATAACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 479
DB 362 GCTGTTTGAACACGGGAAGCATTTATCAGATAAATCAGCAGTGGGAACGAGACCTACCTAG 421
QY 480 GCAATGGTGGTGTGTTGTTGTTATGAGGAGAGCCAGAGTTTAACTGCGAAAGTAAAC 539
DB 422 GCAACGCGCTGGTGTGTTGTTGTTATGAGGAGAGCAGAGTTTAACTGCGAGGCAAGC 481
QY 540 CTGAAGCTGAAGAGAGACTTGTCTTGAAGAATACACTGGGAACACTTACCGAGTGGGTGACA 599
DB 482 CTGAACCTGAAGAGACCTGTTTTCACAAATACACTGGAAACACTTACAAAGTGGGTGACA 541
QY 600 CTTATGACGCTTAAAGACTCCATGATCTGGGACTGTACCTGCTACCTCGGGCTGGGGAG 659
DB 542 CTTATGACGCGCTTAAAGATTCATGATCTGGGACTGTACCTGCAITGGGGCTGGGGAG 601
QY 660 GGAGAAATAGCTGTACCATCGCAACCGCTGCCATGAAGGGGGTCACTCTCTACAAGATTG 719
DB 602 GCAGGATCAGCTGTACCATTCGAATCGTGGCATGAAGGGGGTCACTCTCTACAAGATTG 661
QY 720 GTGACACCTGGAGGAGACCAATGAGACTGGTGTGTTACATGTTAGATGTGTGTCTTG 779

```

Db 662 CTGCAAGTGGAGAGGCCACATGAGACTGGTGGCTATATGTTGGAGTGTGTTGCTCG 721  
Qy 780 GTAATGGAAAAGAGAAATGGACCTGCAAGCCCATAGCTGAGAAAGTGTGTTGATCATCTG 839  
Db 722 GGAATGGAAAAGAGAAATGGACCTGCAAGCCCAATAGCTGAGAAATGTTTGTATCACGCTG 781  
Qy 840 CTGGGACTTCTCTGTTGTTGGTCCGAGAAACGTGGAGAACCCCTACCAAGGCTGGATGATGG 899  
Db 782 CTGGGACTTCTCTGTTGTTGGTCCGAGAAACGTGGAGAACCCCTACCAAGGCTGGATGATGG 841  
Qy 900 TAGATTGTACTTGGCTGGGAGAGCGCAGCAGCATCTTGCACTTCTAGAAATAGAT 959  
Db 842 TGGACTGTACTTGGCTGGGAGAGCGCAGCAGCATCTTGCACTTCTAGAAATAGAT 901  
Qy 960 GCAAGCATCAGGACACAAAGGACATCTCTATAGAAATGGAGACACCTGGAGCAAGAGATA 1019  
Db 902 GCAATGATCAGGACACCAAGGACGCTCTCAGAAATGGAGACACATGGAGCAAGAGACA 961  
Qy 1020 ATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACCGGCCGAGGAGAGTGGAAAGTGTG 1079  
Db 962 ACAGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACCGGCCGAGGAGAGTGGAAAGTGTG 1021  
Qy 1080 AGAGGCAACCTCTGTGTGAGACCAACATCGAGCGGATCTGGCCCTTCAACCGATGTTGCTG 1139  
Db 1022 AGCGACA--TGTTCTACAGAGTGTCTTCAAGTGTGATCTGGCTCTTCAAGATGTCGGA 1078  
Qy 1140 CAGCTGTTTACCAACCGCAGCTCAACCCGAGCTCTCTCCCTATGGCCACTGTGTACAG 1199  
Db 1079 CAGCTATTTACCAACCGCAGCTCAACCCGAGCTCTCTCCCTATGGCCACTGTGTACAG 1138  
Qy 1200 ACAGTGTGTGTACTCTGTGGGATGCAAGTGGTGAAGACCAAGGAAATAGCAAA 1259  
Db 1139 ACAGCGTGTGTGTACTCTGTGGGATGCAAGTGGTGAAGACCAAGGAAATAGCAAA 1198  
Qy 1260 TGCCTTGCAGCTGCTGGGCAACGAGTCACTGCGCAGAGACAGCTGTACCCAGACTT 1319  
Db 1199 TGCTGTGCACTTGGCTGGGCAACGAGTCACTGCGCAGAGACAGCTGTACCCAGACTT 1258  
Qy 1320 AGGTTGGCAACTCAAAATGGAGGCAATGTGTCTTACCAATTCACCTTACAAATGGCAGGAGCT 1379  
Db 1259 AGGTTGGCAACTCAAAATGGAGGCAATGTGTCTTACCAATTCACCTTACAAATGGCAGGAGCT 1318  
Qy 1380 TCTACTCTGCAACACAGAGGCGACAGGACGACATCTTGGTGCAGACACACTTGA 1439  
Db 1319 TCTACTCTGCAACACAGAGGCGGACAGGACATCTTGGTGTAGCAACAATCTCAA 1378  
Qy 1440 ATTATGAGCAGGACACAGAAATACTCTTCTGACAGACCACTGTTTGGTTTCAAGACTC 1499  
Db 1379 ATTATGACACAGACACAGAAATATTCTTCTGACAGACCACTGTTTGGTTTCAAGACTC 1438  
Qy 1500 GAGGAGGAAATTCCAATGGTGTGCTGCGCACTTCCCTTCTCTATACAAACACCAATTT 1559  
Db 1439 GAGGTTGGGAAATTCCAATGGTGTGCTGCGCACTTCCCTTCTCTATACAAACACCAATTT 1498  
Qy 1560 ACACCTGATTGCACTTCTGAGGCGACAGAGCAACATGAAGTGGTGTGGGACCAACAG 1619  
Db 1499 ACAGCGCATGTACTTCTGAGGCTAGCGGCGACACATGAATGGTGTGGGACCAACAG 1558  
Qy 1620 ACTATGATGCGGACAGAAATGTTGGTGTGCGCCCATCGCTGCGCCACAGAGAAATCTGCA 1679  
Db 1559 ACTAGATGCGGACAGAAATGTTGGTGTGCGCCCATCGCTGCGCCACAGAGAGATCTGCA 1618  
Qy 1680 CAAACCAATGAAGGGGTCTATGACCCGATGGAGATCAGTGGGATGAAGCAGATCAGATGG 1739  
Db 1619 CGAACCAAGAGGGGTCTATGATCCGATGGGACCACTGGGATGAAGCAGATCAGATGG 1678  
Qy 1740 GTCACATGATGAGGTGCACTGTGTGGGAAATGGTGTGGGAAATGCAATGCAATGCTCT 1799  
Db 1679 GGCACATGATGAGGTGCACTGTGTGGGAAATGGTGTGGGAAATGCAATGCAATGCTCT 1738  
Qy 1800 ACTCGAGGTTGAGATCAGTGCATTTGATGATCAGTGTGAGGAGGAGGAGGAGGAGGAG 1859

Db 1739 ACTCCAGCTCCGAGATCAGTGCATCGTGTGATGACATTTACTTACAAAGTCAACGACACGT 1798  
Qy 1860 TCCAAAGGCTCATGAAAGAGGGGACATCTGTGAATCTGTACATGCTTCCGTACGGTCCGG 1919  
Db 1799 TCCAAAGGCTCACGAGGAGGACATATGCTGAATCTGTACCTGCTTCGGTCAGGGCCGGG 1858  
Qy 1920 GCAAGGTGGAAAGTGTGATCCCGTCGACCAATGCGAGGATTCAGAGACTGGGAGCTTTTATC 1979  
Db 1859 GAGATGGAAATGTGACCCCATCGACCGATGCCAAGATTCAGAGACCCGACATTTTACC 1918  
Qy 1980 AAATGGAGATTCATGGGAGAAATGTGTCATGTGTGATGATACAGTGTCTACTGCTATG 2039  
Db 1919 AGATTGCTGACTCCTGGGAGAAATTTGTGTCATGTGTGATGATACAGTGTCTACTGCTAG 1978  
Qy 2040 GCGTGGCATTCGGGAGTGGCATTCGCAACCTTTACAGACCTTATCAAGCTCAAGTGTGTC 2099  
Db 1979 GCGTGGCATTCGGGAGTGGCATTCGCAACCTTCTGAGACCTTACCCAGGACAACTGGAC 2038  
Qy 2100 CTGTGCAAGTATTTATCTACTGAGATCTCGAGTCAAGCCCACTCCCAACCCCATCCAGTGA 2159  
Db 2039 CTGTTCAGTAAATTTATCAGGAGACCCCGCAGCCCAATTTCCACCCCATCCAGTGA 2098  
Qy 2160 ATGCACCAAGCATCTCATATTTCCAAAGTACATTTCTCAGGTGGAGACCTAAAAATCTG 2219  
Db 2099 ATGCCCCGAGCTTTCACACATCAACAGTACATTTCTCAGGTGGAGACCTAAAACTCTA 2158  
Qy 2220 TAGGCGTGGGAGGAGCTACCATACCAAGGCCACTTAACTCTCTACACCATCAAGGCC 2279  
Db 2159 CGGTCGCTGGGAGGAGCTACCATTCAGGCGCACTTAACTCTCTATACCATCAAGGCC 2218  
Qy 2280 TGNAGCTGCTGTGTATACGAGGGCCAGCTCATCAGCATCCAGCATCCGCGCCACCAAG 2339  
Db 2219 TGACCCCGAGTGTGATCTACGAGGAGCAGCTCATCAGCATCCAGCATCCGCGCCACCAAG 2278  
Qy 2340 AAGTGACTGCTTTGACTTTCACCAACCAAGGCCACTTAACTCTCTACACCATCAAGGCC 2399  
Db 2279 AAGTGACTGCTTTGACTTTCACCAACCAAGGCCACTTAACTCTCTACACCATCAAGGCC 2335  
Qy 2400 TGACAGGAGAGAGCTCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2459  
Db 2336 TGACTGGAGAGCTCGGCGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2395  
Qy 2460 TCACAGCAGTGTGTGTCT 2519  
Db 2396 TCACAGCAGAGCTTCTGTGTCT 2455  
Qy 2520 GGTTGAAATAGCTGTGATGAGGAGGAGATGAGCCAAGTACCTGTGATCTTCCAGCA 2579  
Db 2456 GAGTGGAGTACGAATCTGAGCGAGGAGATGAGCCTCAGTACCTTGTATCTTCCAGCA 2515  
Qy 2580 CAGCCACTTCTGTGAACATCCCTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2639  
Db 2516 CAGCCACTTCTGTGAACATCCCTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2575  
Qy 2640 ATCAGATATCTGAGGATGGGAGGAGATTTGATCTGTCTACTTCTCTCTCTCTCTCTCTCT 2699  
Db 2576 ATCAGATATCTGAGGAGGAGGAGATTTGATCTGTCTACTTCTCTCTCTCTCTCTCTCTCT 2635  
Qy 2700 CTGATGCGCTCTGTGACCGGAGCTGTGAGCCAGTGTGATGACACCTCAATTTGTTTCTGCT 2759  
Db 2636 CTGATGCGCTCTGTGACCGGAGCTGTGAGCCAGTGTGATGACACTTCCATTTGTTTCTGAT 2695  
Qy 2760 GGAGCAGAGCCCAAGGCTCCCATCAGAGGGGTACAGAAATAGTCTATTTCGCCATCAGTAGAAG 2819  
Db 2696 GGAGCAGAGCCCAAGGCACTTATCAGAGGGGTACAGAAATAGTCTATTTCGCCATCAGTAGAAG 2755  
Qy 2820 GTAGCAGCAGAGACTCAACCTTCTGAACTGTGAACTCCGTCACCTCCCTCAGTGTCTGCT 2879  
Db 2756 GCAGTAGCAGAGACTCAACCTTCTGAACTGTGAACTCCGTCACCTCCCTCAGTGTCTGCT 2815  
Qy 2880 AACCTGCTTCTAGTATAACATCACTTATCTATGCTGTGGAAGAAATCAAGAAATGATACAC 2939  
Db 2816 AGCCCGGTTTCTAGTATAACATCACTTATCTATGCTGTGGAAGAAATCAAGAAATGATACAC 2875

Y	2940	CTGTTGCTATTCAACAGAGAAACCACTGGGACCCACCGCTCAGATACAGTGCCTCTCCCA	2999	QY	4020	AGGATGACAGGAAAGTGTCCTTATCTCTGATACCATCATCCCA	4063
b	2876	CCGTTTTCATCCAGCAGGAGACTCTGGCGTCCACGATCCGATGATGTCCCGTCCAA	2935	Db	3956	AGATGACAGGAAAGTGTCCTTATCTCTGATACCGTCACTCCAGAGTGCCTCCAGCTCA	4015
Y	3000	GGGACCTGACAGTTTGTGGAAGTGACAGAGCTGAAGGTACCATCATGTGGACACCGCTG	3059	QY	4064	-----	4063
b	2936	AGAACCTACAGTTTGTGGAAGTGACCGAGCTGAAGTACCATCATGTGGACACCTCCTA	2995	Db	4016	CTGACCTAAGCTTTTGTGATATACTGACTCAAGCATCGGCCTGAGGTGGACCCCGCTAA	4075
Y	3060	AGAGTGACATGACCGGCTACCGTGTGATGTGATCCCGTCAACCTGCTGGCGAGACAG	3119	QY	4064	-----	4063
b	2996	ATAGCGACGTGACTGGATACCGTGTGATGTCTGCTGTCAACCTGCCAGGGGAACATG	3055	Db	4076	ACTCTCCACCATATCGGATACCGAATCACAGTAGTTGGGCGCAGAGAGGGATCCCCA	4135
Y	3120	GGCAGAGGTCGCCATCAGCAGGAACACCTTTGAGAGTCAACCGGCTGTCCCTGGGG	3179	QY	4064	-----	4063
b	3056	GGCAGAGGCTGCTGTCAACAGGAACCTTTGCTGAAGTCAACCGGACTGTCCCGAGGG	3115	Db	4136	TTTTTGAAGATTTTGTGACTCCTCAGTAGGATACTACAGTTACAGGGCTGGAAACCG	4195
Y	3180	TCACTTATTACTTCAAAAGTCTTTGAGTGAGCCATGGGAGGAGAGCAAGCCTCTGACTG	3239	QY	4064	-----	4063
b	3116	TCACTTATTACTTCAAAAGTCTTTGAGTGAGCCATGGGAGGAGAGCAAGCCTCTGACTG	3175	Db	4196	GCATGTGACTATGACATCAGCGTTATCCTCTCATTAATGGCGGAGAGAGTGCCCTACTA	4255
Y	3240	CTCAACAGACACCAACTGGAATGCTCCCACTAACCTCCAGTTTGTCAATGAATCTGATT	3299	QY	4064	-----	4106
b	3176	CACAAAGACACCAACTGGAATGCTCCCACTAACCTCCAGTTTGTCAATGAATCTGATT	3235	Db	4256	CACGTACACAGCAAAACGGCGTCCCTCCTCCACGGATCTGCGATTTCACCAATATCGGTC	4315
Y	3300	CTACTGTCTGTGATGGAATCTCACTCGGCGCCAGATACAGGATACCGACTGACCG	3359	QY	4107	CAGACACATCGGTGTCACTGGGCTCCACCCCATCCATTGATTTAAACCACTCTCTGG	4166
b	3236	GAACAGTTCTGGTAACTTGGACTCCACCTCGAGCCCGGATAGCAGGCTACCGACTGACAG	3295	Db	4316	CGGACACTATGCGGGTCACTTGGGCGCCGCTCCGTCATTGAGCTAACCACTCTTGG	4375
Y	3360	TGGGCTTACCCGAAAGAGACAGCCAGCAGTACAAATGTGGTCCCTCTGTCCAGT	3419	QY	4167	TGCGTTACTCACTGTGAAAAATGAGGAAGATGTTCAGAGTTGTCAATTTCTCTTTCAG	4226
b	3296	TGGGCTTACCCGAAAGAGACAGCCAGCAGTACAAATGTGGGACCCATGTGGTCCAGT	3355	Db	4376	TGCGTACTCACTGTGAAAGACGAGGAGATGTGGCAGAGCTGTCCATTTCACTCCCTCAG	4435
Y	3420	ACCCACTGAGGAATCTGACGCTGCTGCTGAGTACACCGTATCCCTCGTGGCCATAAAGG	3479	QY	4227	ACAATGAGTGGTCTTAAACAAATCTCTGCTCGTACAGATATGTAGTGTGAGTGTCTCCA	4286
b	3356	ATCCCTGAGAAATCTGACGCTGGTCTGAGTACACTGTGACCTTGTATGGTGTGAAG	3415	Db	4436	ACAAACGCGTGGTCTTAAACAAATCTCTGCTCGGACTGAGTACCTAGTGTCTCCA	4495
Y	3480	GCAACAGAGAGCCCAAGCCACTGAGTCTTTACCACTGAGGCTGGAGCTCTA	3539	QY	4287	GTGTCTACGAAACATGAGAGACACCTCTTAGAGGAACAGAGAAACAGAGTCTTGAAT	4346
b	3416	GCAACAGAGAGTCCCAAGCCACCGAGTCTTTACTACCTGAGGCTCTGCGCTCCA	3475	Db	4496	CGGTGTACGAAACATGAGAGACACCTCTCAGAGGAACAGAGAAACAGGCTGGACT	4555
Y	3540	TTCCACTTACAAACCGAGGTGACTGAGACCAACCATTTGTGATCAGATGAGGCGCTGCTC	3599	QY	4347	CCCAACTGAGTGTCTTTGATTTCTCTGATGTACCGCCCACTCAITCACCGTCCACTGGGTGG	4406
b	3476	TTCCACTTACAAACCGAGGTGAGACCAACCATTTGTGATCAGATGAGGCGCTGCTC	3535	Db	4556	CCCAACTGAGTGTCTTTGATTTCTCTGATGTACCGCCCACTCAITCACCGTCCACTGGGTGG	4615
Y	3600	CAAGAAATGGTTTAAAGTGGGTGTAGCAACCAAGCCAGGAGGAGGACCAACAGAGAG	3659	QY	4407	CTCTTCGAGCCACCATCACTGGGTACAGGATCGCCATCATCCGAGCACTTCAGTGGGA	4466
b	3536	CAAGAAATGGTTTAAAGTGGGTGTAGCAACCAAGCCAGGAGGAGGACCAACAGAGAG	3595	Db	4616	CTCTTCGAGCCACCATCACTGGGTACAGGATCGCCATCATCCGAGCACTTCAGTGGGA	4675
Y	3660	TGACTTCAGACTCAGGAAGCATCGTTGTGTCGGCTTGACTCCAGGAGTAGAATACGTCT	3719	QY	4467	GACCTCGAGAAAGATCGGGTGCCTCTCGGAAATTCATCACTCACTCCCTCACCACTCTAATC	4526
b	3596	TGACTTCAGACTCAGGAAGCATCGTTGTGTCGGCTTGACTCCAGGAGTAGAATACGTCT	3655	Db	4676	GACCCAGCAAGACCGAGTGCCTCTCAAGGAATTTCTATCACCTCTCACCACTTAAATC	4735
Y	3720	ACACCATCCAAAGTCTGAGAGATGGAAGAGAGATGGGCCAATTTGAACAAGTGG	3779	QY	4527	CAGGCACAGAGTATGTGCTCAGCATCGTCTCTTATGCGAGAGAGAAAGTCCCTTAT	4586
b	3656	ACACCATCCAAAGTCTGAGAGATGGAAGAGAGATGGGCCAATTTGAACAAGTGG	3715	Db	4736	CGGCACCGAGTACATTTGTACCATCATTTGTGTATGCGAGAGAGAGAGGCCCCAC	4795
Y	3780	TGACACCATGCTCCACCAACAACTTGATCTGGAGGCAACCTGACTGGAGTGC	3839	QY	4587	TGATGGCCCAACATCAACAGTTTCTGATGTTCGAGGACCTGGAAAGTTGTGCTCGGA	4646
b	3716	TGACACCATGCTCCACCAACAACTTGATCTGGAGGCAACCTGACTGGAGTGC	3775	Db	4796	TGATGGCCCAACATCAACAGTTTCTGATGTTCGAGGACCTGGAAAGTTGTGCTCGGA	4855
Y	3840	TCAAGTCTCTCGGAGAGAGACCAACCCAGACATTTACTGTTATAGAAATACACAA	3899	QY	4647	CCCCACCGACTACTGATCAGCTGGGATGCTCTGCTGTCACAGTGCAGATATTACAGGA	4706
b	3776	TTACCGTCTCTCGGAGAGAGACCAACCCAGACATTTACTGGCTACAGAAATACACCA	3835	Db	4856	CCCCACCGACTACTGATCAGCTGGGATGCTCTGCTGTCACAGTGCAGATATTACAGGA	4915
Y	3900	CCCTACAAAGCGCCAGCGGAAATTTCTTGAAGAGTGGTCCATGCTGATCAGACT	3959	QY	4707	TCACCTTACGGAAGAACAGGAGAAATAGCCCTGTCCAGGAGTTTCACTGTGCTGGGAGCA	4766
b	3836	CCCCCAAAAGCGCCAGCGGAAATTTCTTGAAGAGTGGTCCATGCTGATCAGACT	3895	Db	4916	TCACCTTACGGAAGAACAGGAGAAATAGCCCTGTCCAGGAAATTCACCTGTGCCCGGAAGCA	4975
Y	3960	CCTGCACTTTTGAATCACTGAGTCCCGGCTCGAGTACAAATGTCACTGTGTTACCTGCA	4019	QY	4767	AGTCTACAGCTACCATCAGCGGCTTAAACCTGAGTTGATTTACCATCACTGTGTATG	4826
b	3896	CCTGCACTTTTGAATCACTGAGTCCCGGCTCGAGTACAAATGTCACTGTGTTACCTGCA	3955	Db	4976	AGTCTACAGCTACCATCAGCGGCTTAAACCTGAGTTGATTTACCATCACTGTGTATG	5035
				QY	4827	CTGTCACTGGCGGTGGAGACAGCCCGCCAGAGCAAGCAAAATTTCCATTAATTAACGAA	4886

Db 5036 CTGTCACTGGCGTGGGACAGTCCAGCCAGCAGCAAGCCAGTTTCCATCATTAATCAAA 5095  
 Qy 4887 CAGAAATTGACAAAACATCCAGATGCAAGTACCGATGTTCCAGGACAAACAGCATTAGTG 4946  
 Db 5096 CAGAAATTGACAAAGCAATCCAGATGCAAGTACCGATGTTCCAGGACAAACAGCATTAGTG 5155  
 Qy 4947 TCAAGTGGCTGCTTCAAGTCCCTCTGTTACCTGTTACAGAGTAAACCACTCCCAAAA 5006  
 Db 5156 TCAGGTGGCTGCTTCAAGTCCCTCTGTTACAGAGTAAACCACTCCCAAAA 5215  
 Qy 5007 ATGGACAGGACCAACAAAACATTAACCTGAGGTCCAGATCAACAGAAATGACTATTG 5066  
 Db 5216 ATGGCTAGGACCAACAAAATCTAAATCTGAGTCCAGATCAACAGAAATGACTATTG 5275  
 Qy 5067 AAGGCTTGAGCCCAAGTGGAGTATGTTGTTAGTGTCTATGCTCAGAAATCCAGAGGAG 5126  
 Db 5276 AAGGTTGCAACCCACCGTGGAGTATGTTGTTAGTGTCTATGCTCAGAAACCGGACGAG 5335  
 Qy 5127 AGAGTCAGGCTCTGGTTGAGCTGAGTAAACCAACATGATGCGCTTAAGAGCTGGCAT 5186  
 Db 5336 AAAGCCAGCCCTGGTTGAGCTGAGTAAACCAACATGATGCGCTTAAGAGCTGGCAT 5395  
 Qy 5187 TCAGTGTGATGATGTCAGTCCATCAAAATTTGCTTGGAAAGCCCAAGGCGCAAGTTT 5246  
 Db 5396 TCAGTGTGATGATGTCAGTCCATCAAAATTTGCTTGGAAAGCCCAAGGCGCAAGTTT 5455  
 Qy 5247 CCAGGTACAGGCTGCTGACTGAGCCCTGAGATGGAATCCATGAGTATTTCCCTGAC 5306  
 Db 5456 CCAGGTACAGGCTGCTGACTGAGCCCTGAGATGGAATCCATGAGTATTTCCCTGAC 5515  
 Qy 5307 CTGATGTTGAGAAACATGTCAGAGCTGCAAGGCTCAGACCGGTTCTGAGTACACAG 5366  
 Db 5516 CTGATGTTGAGAAACATGTCAGAGCTGCAAGGCTCAGACCGGTTCTGAGTACACAG 5575  
 Qy 5367 TCAGTGTGTTGCTTGGACGATGATGGAGCCAGCCCTGATGGAAACCCAGTCCA 5426  
 Db 5576 TCAGTGTGTTGCTTGGACGATGATGGAGCCAGCCCTGATGGAGTCCAGTCCA 5635  
 Qy 5427 CAGCTATTCTGCAACCAATGTTCACTGAGTATCGAGTGGCGGTGACCCCAAGG 5546  
 Db 5636 CAGCCATTCTGCGCAACCAATGTTCACTGAGTATCGAGTGGCGGTGACCCCAAGG 5695  
 Qy 5487 CCCAGTGGACACCAACCAATGTTCACTGAGTATCGAGTGGCGGTGACCCCAAGG 5546  
 Db 5696 CCCAGTGGACACCGCCAGTGTAGCTCACTGAGTATCGAGTGGCGGTGACCCCAAGG 5755  
 Qy 5547 AGAAGCCGACCAATGAAAGAAATCAACCTTGTCTTCAAGCTCATCCGTGGTTGAT 5606  
 Db 5756 AGAAGACAGGACCAATGAAAGAAATCAACCTTGTCTTCAAGCTCATCCGTGGTTGAT 5815  
 Qy 5607 CAGGACTTATGGTGGCCACCAATATGAAGTGGTGTCTATGCTCTTAAGGACACTTTGA 5666  
 Db 5816 CAGGCTCATGGTGGCCACCAATATGAAGTGGTGTCTATGCTCTTAAGGACACTTTGA 5875  
 Qy 5667 CAGGACAGCTCAGGAGTTGTCAACCTTGGAGATGTCAAGCCCAACAGAGGG 5726  
 Db 5876 CAGGACAGCTCAGGAGTTGTCAACCTTGGAGATGTCAAGCCCAACAGAGGG 5935  
 Qy 5727 CTCGTGTGACAGATGCTACTGAGACCAACCATCAGTATAGCTGGAGAACCAAGCTGAGA 5786  
 Db 5936 CCGGTGTGACAGATGCTACTGAGACCAACCATCAGTATAGCTGGAGAACCAAGCTGAGA 5995  
 Qy 5787 CGATCACTGGCTCAAGTTGATGCGTTCCAGCCAAATGGCCAGACTCCAAATCCAGAGAA 5846  
 Db 5996 CGATCACTGGCTTCAAGTTGATGCGTTCCAGCCAAATGGCCAGACTCCAAATCCAGAGAA 6055  
 Qy 5847 CCATCAAGCCAGATGTCAGAGCTACCATCAGTCTTACACAGGCGCTGACTACA 5906  
 Db 6056 CCATCAAGCCAGATGTCAGAGCTTATACTATAGGTTTACAGCCGCTGACTACA 6115  
 Qy 5907 AGATCTACTGTCAGCTTGAATGATGCTCGGAGCTCCCTGTGTGCTATCGAGCCT 5966

6116 AGATCCACCTGTACAGCTCAACGACAAATGCCGGAGCTCTCTGTGTGTCATTGATGCCT 6175  
 Qy 5967 CCACTGCCATTGATGACACCATCAACCTGCTTCTGTCGCCACCAACCAATTCCTTCCG 6026  
 Db 6176 CCAGGCCATTGATGCCCATCAACCTGCTTCTGACCAACCAACCAATTCCTTCCG 6235  
 Qy 6027 TGTATCATGTCAGCCGCGCACGTCGCCAGGATTACCGCTTACATCATCAAGTATGAGAAC 6086  
 Db 6236 TGTATCATGTCAGCCGCGCACGTCGCCAGGATTACCGCTTACATCATCAAGTATGAGAAC 6295  
 Qy 6087 CTGGTCTCTCCACAGAGTGGTCCCTGCGCCCGCTGCTGTCACAGAGGCTACTTA 6146  
 Db 6296 CTGGTCTCTCCACAGAGTGGTCCCTGCGCCCGCTGCTGTCACAGAGGCTACTTA 6355  
 Qy 6147 TTAATGCTTGGAAACCGGAAACCGAATATACAAATTTATGTCATTGCTTGAAGAAATTAATC 6206  
 Db 6356 TCACTGCTTGGAGCCAGAACCGAGTACACCATCTATGTCATCGCACTGAAGAACATC 6415  
 Qy 6207 AGAAGAGCGAGCCCTGATTTGGAAGAAAAGACAGACAGCTTCCCACTGGTTAACCC 6266  
 Db 6416 AGAAGAGTGGCCCTGATTTGGAGGAAAAGACAGATGAGCTTCCCACTGGTTAACCC 6475  
 Qy 6267 TTCCACACCCCAATCTTCAATGACACAGATCTTGGATGTTCTTCCACAGTTCAAAAGA 6326  
 Db 6476 TTCCACACCCCAATCTTCAATGACACAGATCTTGGATGTTCTTCCACAGTTCAAAAGA 6535  
 Qy 6327 CCCTTTGTCACCCACCTCGGTATGACACTGGAATGTTATGAGTTCTCTGGCACTT 6386  
 Db 6536 CCCTTTGTCACCCACCTCGGTATGACACGGAATGTTATGAGTTCTCTGGCACTT 6595  
 Qy 6387 CTGGTCAAGCAACCCAGTGTGGCAACAAATGATCTTTCAGGAAACATGTTTTCAGCGGA 6446  
 Db 6596 CCACACCAACCCAGTGTGGCAACAAATGATCTTTCAGGAAACATGTTTTCAGCGGA 6655  
 Qy 6447 CCACACCGCCCAACCGGCAACCCCATAGGCATAGGCCAAGACCATACCCCGCAATG 6506  
 Db 6656 CCACGCAACCACTGCGGCCACCCCGTCAGGCTTAGGCCAAGACCATACCTGCGCAATG 6715  
 Qy 6507 T----- 6507  
 Db 6716 TAGATGAGGAGTCCAAATCGGTCTATGTTCCAGGGGAGACGTAGACTACCACTCTATC 6775  
 Qy 6508 -----AGGCAAGAAAGCTCTCTCTCAGACAA 6533  
 Db 6776 CTCATGTTCCGGGCTCAATCCAAATGCTCTACAGGACAAAGAGCTCTCTCTCAGACAA 6835  
 Qy 6534 CCACTCATGGGCCCATTTCCAGGACACTCTGAGTACATCATTTTCATGTCATCTGTTG 6593  
 Db 6836 CCACTCTTGGAGCCCATTTCCAGGAGTCTGAGTACATCATTTTCATGTCATCTGTTG 6895  
 Qy 6594 GCATGATGAAGAACCCCTTACAGTTCAGGTTCCCTGGAACCTTCTACAGTGCACCTCTGA 6653  
 Db 6896 GCATGACGAAGAGCCCTTACAGTTCGAAGTTCCTGGAACCTTCTACAGTGCACCTCTGA 6955  
 Qy 6654 CAGGCTCACCAGAGTGCACCTACAAATCATAGTGGAGGACCTGAAGACCAAGCAGAGA 6713  
 Db 6956 CTGGCTTACAGAGGGTCACTACAAATCATAGTGGAGGCTTGCACAAACCAAGAGA 7015  
 Qy 6714 GGCATAAGGTTCCGGAAGAGTGTTCCTGCGGCAACTCTGTCAACGAAGGCTTGAAC 6773  
 Db 7016 GGCATAAGTCCGGAAGAGTGTTCCTGCGGCAACTCTGTCAACGAAGGCTTGAAC 7075  
 Qy 6774 AACCTACGAGTACTGCTGTTGACCCCTACAGTGTCCCATATGCGTGTGGAGATG 6833  
 Db 7076 AGCTACGAGTACTGCTGTTGACCCCTTACAGGTTTCCCATACGCGTGTGGAGAG 7135  
 Qy 6834 AGTGGGAAACGAATGCTGGAATCAGGCTTTTAAACTGTTGCGCAGTGTCTTAGCTTTGGA 6893  
 Db 7136 AATGGAGCGGTTATCTGACTCTGGCTTTAGCTCACTTGGCAGTGTCTGGGCTTTGGA 7195  
 Qy 6894 GTGGTCACTTCAGATGATGATCTAGATGTTGCGATGACAAATGTTGAACTACAAAGA 6953  
 Db 7196 GTGGTCACTTCAGATGATGATCTAGATGTTGCGATGACAAACGTTGCAACTACAAAGA 7255

6954 TTGGAGAGAGTGGACCGTCAAGGAGAAATGGCCAGATGATGAGCTGCACATGTCTTG 7013  
b |||||  
7256 TCGGAGAGAGTGGGATCGTCAAGGAGAAATGGCCAGGATGAGCTGCACATGTCTCG 7315  
y |||||  
7014 GGAACGGAAAGGAGAAATCAAGTGTGACCTCATGAGGCAACGTGTTPACGATGAGGA 7073  
b |||||  
7316 GGAATGGAAGGGAGAAATCAAAATCGGATCCCATGAAGCAACGTGTATGACGACGGA 7375  
y |||||  
7074 AGACATACACATGAGAGACAGTGGCAGAGGATATCTGGTGCCATTTGCTCCTGCA 7133  
b |||||  
7376 AGACCTACACATGAGAGACAGTGGCAGAGGATATCTGGAGCCATTTGCTCCTGCA 7435  
y |||||  
7134 CATGCTTTGGAGGCCAGCGGGCTGGCGTGTGACAACCTGCCGACAGACCTGG--GGTG 7190  
b |||||  
7436 CGTGTTCGGGGGCCAGCGGGCTGGCGTGTGACAACCTGCCGACAGACCTGGGGCTGCTG 7495  
y |||||  
7191 AACCCAGTCCGAGGACATGAGTGGCCAGTCTGACCAACAGTATTTCTCAGAGATACCATC 7250  
b |||||  
7496 AACCCAGTCCGAGTACCTGACCTGGCCACACCTACACAGTATACAGAGATACCATC 7555  
y |||||  
7251 AGAGAACAAACACTAATGTTAAATGGCCAAATGAGTGTCTTCAATGCTTTAGATGTACAGG 7310  
b |||||  
7556 AGGAAAGHACATATGTAATGGCCAAATGGAATGCTTCAATGCGTTGGACGTGCAGG 7615  
y |||||  
7311 CTGACAGAGAGATCCCGAGAGTAAATCATCTTTCCATCCAGAGGACAAAGCATGTCT 7370  
b |||||  
7616 CTGACAGAGATATCCAGAGATATCTTCCATCCAGCCCAAGCCCAAG--TGCT 7673  
y |||||  
7371 CTCTGCCAAGATCCATCTAACTGGAGTGAATGTAGCAGAC---CCAGCTTAGAGTTCTT 7427  
b |||||  
7674 CTCTACCAAGTCAATCCACACCCAGTGAATGTAGCAGACCTCCATTTCTGAGTGTG 7733  
y |||||  
7428 CTCTCTTTCTTAAGCCCTTGTCTGAGAGAGTCTCCAGCTTCAGCTCAACTCACAGC 7487  
b |||||  
7734 ATTTCAACCTTAAGCCTTCTGCTCTGAGTCAAGTTCTCAGCTTCAGCTCAACTTACAGC 7793  
y |||||  
7488 TTCTCCAAGCATCACCTG-GGAGTTTCTGAGGGTTTCTCATAAATGAGGGCTGCACA 7546  
b |||||  
7794 TTCTCCAAGCATCCCGCGGGATGTTTGAGACTTCCCTCTTAAATGGTGACAGTTGG 7853  
y |||||  
7547 TTGCTGTTCTGTTGAAAGTATTCATACCGCTCAGTATTTTAAATGAAGTGAATCTAA 7606  
b |||||  
7854 TGCCCTGTTCTGCTTCAGGGTATTCAGTACTGCTCAGTATTTATGTTCTAAGAGAAATCAA 7913  
y |||||  
7607 -----GATTTGGTTGGGATCAATGGAAGCATATGAGCCCAACCAAGATGCAAA 7657  
b |||||  
7914 AGTCTGTGATTGGTCTGGATCAA-AGGAAACACAGGTAGCCACACGATGC-AA 7971  
y |||||  
7658 TGTTTGAATGATGACCAAAATTTTAAGTAGGAAAGTCAACCAAACTTCTGCTTT 7717  
b |||||  
7972 TGAATTGAATGGTAGTACCCCAAGAGGGGAGCAGGAAGTTAAACACAGACAGTTCTGCTTT 8031  
y |||||  
7718 CACTT 7722  
b |||||  
8032 CTTT 8036

Search completed: March 8, 2004, 04:36:35  
Job time : 2579 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:19:42 ; Search time 69.2294 seconds  
(without alignments)

12152.440 Million cell updates/sec

Title: US-10-084-817-8

Perfect score: 1516

Sequence: 1 ctttgttttggacatagct.....gctgaaataaaatgncgtc 1516

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA.\*
- 1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1515	99.9	1516	4	US-09-919-172-38
2	1387.8	91.5	1399	4	US-09-023-655-1342
3	332.8	22.0	1001	4	US-09-641-638-86
4	319.4	21.1	1362	4	US-09-023-655-1152
5	268.4	17.7	1339	4	US-09-023-655-1149
6	250.2	16.5	2305	3	US-08-526-136-1
7	250.2	16.5	2311	3	US-08-526-136-3
8	245.2	16.2	1301	4	US-09-023-655-1190
9	239.8	15.8	957	1	US-08-125-746-4
10	239.8	15.8	1460	6	5225537-3
11	239.8	15.8	1567	1	US-08-125-746-2
12	236.6	15.6	981	4	US-09-324-096A-1
13	236.6	15.6	981	4	US-09-324-096A-3
14	236.6	15.6	981	4	US-09-324-096A-5
15	232.8	15.4	2137	4	US-09-976-594-348
16	216	14.2	7938	3	US-09-331-581-14
17	215.2	14.2	7286	3	US-09-331-581-3
18	184.6	12.4	1356	4	US-09-010-147B-15
19	184.6	12.4	1356	4	US-09-641-638-80
20	158	10.4	1793	4	US-09-620-312D-747
21	153.2	10.1	644	3	US-09-328-111-721
22	141.2	9.3	502	3	US-09-328-111-625
23	141.2	9.3	630	3	US-09-328-111-491
24	135.4	8.9	501	3	US-09-328-111-481
25	125.4	8.2	1293	4	US-09-325-932A-17
26	124.4	8.2	1001	4	US-09-641-638-106
27	124.4	8.2	1001	4	US-09-641-638-107

C	28	118.8	7.8	643	3	US-09-385-982-306	Sequence 306, App
	29	116.6	7.7	741	4	US-09-641-638-81	Sequence 81, Appl
	30	102.2	6.7	543	3	US-09-328-111-428	Sequence 428, App
	31	101.2	6.7	616	4	US-09-621-976-1120	Sequence 1120, Ap
	32	101.2	6.7	813	4	US-09-641-638-100	Sequence 100, App
	33	101.2	6.7	943	4	US-09-641-638-39	Sequence 99, Appl
	34	100.6	6.6	396	4	US-09-640-173-60	Sequence 60, Appl
	35	100.6	6.6	396	4	US-09-713-550-60	Sequence 60, Appl
	36	99.2	6.5	789	4	US-09-325-932A-20	Sequence 20, Appl
	37	95.4	6.3	903	4	US-09-641-638-83	Sequence 83, Appl
	38	83.2	5.5	1001	4	US-09-641-638-115	Sequence 115, App
	39	82.4	5.4	1409	4	US-09-976-594-349	Sequence 349, App
	40	72.2	4.8	442	4	US-09-621-976-16600	Sequence 16600, A
	41	70	4.6	401	4	US-09-643-597-261	Sequence 261, App
	42	70	4.6	401	4	US-09-480-884A-261	Sequence 261, App
	43	70	4.6	401	4	US-09-542-615A-261	Sequence 261, App
	44	70	4.6	401	4	US-09-606-421B-261	Sequence 261, App
	45	61.4	4.1	1000	4	US-09-641-638-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1  
US-09-919-172-38

; Sequence 38, Application US/09919172

; Patent No. 6673545

; GENERAL INFORMATION:

; APPLICANT: Paris, Mary

; APPLICANT: Turner, Christopher M.

; TITLE OF INVENTION: PROSTATE CANCER MARKERS

; FILE REFERENCE: PA-0036 US

; CURRENT APPLICATION NUMBER: US/09/919,172

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/222,469

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 102

; SOFTWARE: PERL Program

; SEQ ID NO 38

; LENGTH: 1516

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6673545 1303785CB1

; NAME/KEY: unsure

; LOCATION: 1512

; OTHER INFORMATION: a, t, c, g, or other

; US-09-919-172-38

Query Match 99.9%; Score 1515; DB 4; Length 1516;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CTTTGTTTTGGACATAGCTGAGCCATGTACTTCAACAGAGGCGAGCCAAATTAATACT	60
Db	1	CTTTGTTTTGGACATAGCTGAGCCATGTACTTCAACAGAGGCGAGCCAAATTAATACT	60
Qy	61	TCGTGTTGCTAGTGTGGCTTCCTTTAAATCTATATAATCAGAGCCCAAGTCTCCAC	120
Db	61	TCGTGTTGCTAGTGTGGCTTCCTTTAAATCTATATAATCAGAGCCCAAGTCTCCAC	120
Qy	121	TGCCAGTGTGAATCTTCAGAGAGAAATCTCTTTAGTCTTTTTCGCAAGAAGTACAGAT	180
Db	121	TGCCAGTGTGAATCTTCAGAGAGAAATCTCTTTAGTCTTTTTCGCAAGAAGTACAGAT	180
Qy	181	AAAGACACTTTTTCAAAAATGGAATGGTATCAGAAATCTCTCAAGCAGCCCTGTTTATT	240
Db	181	AAAGACACTTTTTCAAAAATGGAATGGTATCAGAAATCTCTCAAGCAGCCCTGTTTATT	240
Qy	241	GAATAATGAGAGCAGGAATATGTTCAACTGTGAAGTCTCAAGGTTGTCCTCCGATCA	300
Db	241	GAATAATGAGAGCAGGAATATGTTCAACTGTGAAGTCTCAAGGTTGTCCTCCGATCA	300

QY	301	GGGGTGAGGCCCTATCCTTACCTTCCTCCGATGTGCCTTGCATAAGGCC	360
DB	301	GGGGTGAGGCCCTATCCTTACCCTCAATCCATCCTCGATGTGCCTTGCATAAGGCC	360
QY	361	ATAATGGTTAAAGGTGTGATGAAGCAACCATTCATTCTTAACCTAAGCGAACCAAT	420
DB	361	ATAATGGTTAAAGGTGTGATGAAGCAACCATTCATTCTTAACCTAAGCGAACCAAT	420
QY	421	GCAACGCTCAACAGATCAAAGCAGCATATCTCCAGGAAACAGGAAAGCCCCTGGATGAA	480
DB	421	GCAACGCTCAACAGATCAAAGCAGCATATCTCCAGGAAACAGGAAAGCCCCTGGATGAA	480
QY	481	ACACTGAAGAAGGCCCTTACAGGTCACTTGAAGGTGTGTTTTAGCTCTGCTAAAACT	540
DB	481	ACACTGAAGAAGGCCCTTACAGGTCACTTGAAGGTGTGTTTTAGCTCTGCTAAAACT	540
QY	541	CCAGCGCAATTTGATGCTGATGAACCTTCGTGTCATGAAGGCGCTTGGAACTGATGAA	600
DB	541	CCAGCGCAATTTGATGCTGATGAACCTTCGTGTCATGAAGGCGCTTGGAACTGATGAA	600
QY	601	GATACTCTAATTCAGATTTTGGCATCAAGACTCAACAAGAAATCAGAGACATTAACAGG	660
DB	601	GATACTCTAATTCAGATTTTGGCATCAAGACTCAACAAGAAATCAGAGACATTAACAGG	660
QY	661	GTCTACAGAGAGAACTGAAGAGAGATCTGGCCAAAAGACATAACCTCAGACATCTGGA	720
DB	661	GTCTACAGAGAGAACTGAAGAGAGATCTGGCCAAAAGACATAACCTCAGACATCTGGA	720
QY	721	GATTTTCGGNAACGCTTGTCTTCTCTGCTTAAGGTGACCGATCTGAGACATTCGTGTG	780
DB	721	GATTTTCGGNAACGCTTGTCTTCTCTGCTTAAGGTGACCGATCTGAGACATTCGTGTG	780
QY	781	AATGAAGACTTGCTGATTCAGATGCGCAGGCGCTTGTATGAAGCAGGAGAAAGGAAAG	840
DB	781	AATGAAGACTTGCTGATTCAGATGCGCAGGCGCTTGTATGAAGCAGGAGAAAGGAAAG	840
QY	841	GGACAGACGTAAACGTGTCTCAATACCATCCTTACCACAGAGCTATCCAACATTCGC	900
DB	841	GGACAGACGTAAACGTGTCTCAATACCATCCTTACCACAGAGCTATCCAACATTCGC	900
QY	901	AGAGTGTTCAGAAAAATACACCAAAGTACAGTAAGCATGACATGAACAAAGTTCTGGA	960
DB	901	AGAGTGTTCAGAAAAATACACCAAAGTACAGTAAGCATGACATGAACAAAGTTCTGGA	960
QY	961	GAGTTGAAGGTGACATGTGAATAATGCTCACAGCTATCGTGAAGTGCAGCAAGCAAA	1020
DB	961	GAGTTGAAGGTGACATGTGAATAATGCTCACAGCTATCGTGAAGTGCAGCAAGCAAA	1020
QY	1021	CCAGCTTTCTTTGCAGAGAAAGCTTCATCAAGCCATGAAAGGTGTGGAACTCGCCATAAG	1080
DB	1021	CCAGCTTTCTTTGCAGAGAAAGCTTCATCAAGCCATGAAAGGTGTGGAACTCGCCATAAG	1080
QY	1081	GCATTGATCAGGATTATGGTTCCCGTTCGAAATTGACATGATGATCAAGCATTC	1140
DB	1081	GCATTGATCAGGATTATGGTTCCCGTTCGAAATTGACATGATGATCAAGCATTC	1140
QY	1141	TATCAGAAGATGTATGGTATCTCCCTTTGCCAAGCCATCCTGGATGAAACCAAGGAGAG	1200
DB	1141	TATCAGAAGATGTATGGTATCTCCCTTTGCCAAGCCATCCTGGATGAAACCAAGGAGAG	1200
QY	1201	TATGAGAAAATCCTGGTGGCTCTTTGTGGAGGAACCTAAACATTCCTTGATGGTCTCAA	1260
DB	1201	TATGAGAAAATCCTGGTGGCTCTTTGTGGAGGAACCTAAACATTCCTTGATGGTCTCAA	1260
QY	1261	GCTATGATCAGAAGACTTTAATTATATATTTTCA TCCTATAAGCTTAAATAGGAAAGTTT	1320
DB	1261	GCTATGATCAGAAGACTTTAATTATATATTTTCA TCCTATAAGCTTAAATAGGAAAGTTT	1320
QY	1321	CTTCAACAGGATTACAGTGTAGCTACCTACATGCTGTAATAATATAGCCTTTAATCATTT	1380
DB	1321	CTTCAACAGGATTACAGTGTAGCTACCTACATGCTGTAATAATATAGCCTTTAATCATTT	1380

```

QY      1381 TTATATTATAACTCTGTATTAATAGAGATAAGTCCCAATTTTAAATAAATGTTTTCCCCCAAAC   1440
DB      1381 TTATATTATAACTCTGTATTAATAGAGATAAGTCCCAATTTTAAATAAATGTTTTCCCCCAAAC   1440
QY      1441 CATAAAACCCTATACAAGTGTTCTTAGTAACAATACATGAGAAAAATGCTCTATGTAGCTG   1500
DB      1441 CATAAAACCCTATACAAGTGTTCTTAGTAACAATACATGAGAAAAATGCTCTATGTAGCTG   1500
QY      1501 AAAATAAAATGNCGTC 1516
DB      1501 AAAATAAAATGNCGTC 1516

RESULT 2
US-09-023-655-1342
; Sequence 1342, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocke, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1342:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1399 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g34387
; US-09-023-655-1342

```

	Query Match	91.5%	Score 1387.8	DB 4	Length 1399	
	Best Local Similarity	99.8%	Pred. No. 0			
	Matches 1389	Conservative	0	Mismatches 3	Indels 0	Gaps 0
Qy	125	AGTGTGAATCTTCAGAGAAGAA	TTCTCTTTAGTCTTTTCGGAAGAGGTAGAGATAAG	184		
Db	1	AGTGTGAATCTTCAGAGAAGAA	TTCTTTAGTCTTTTCGGAAGAGGTAGAGATAAG	60		
Qy	185	ACACTTTTTCAAAAATGGCAAT	TGTTATCAGAAATTCCTCAAGCAGGCCTGTTTTATTGAAA	244		
Db	61	ACACTTTTTCAAAAATGGCAAT	TGTTATCAGAAATTCCTCAAGCAGGCCTGTTTTATTGAAA	120		

```

245 ATGAAGAGCGGAATATGTTCAACTGTGAAGTCATCCAAAGGTGTCTCCGGATCAGCGG 304
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 ATGAAGAGCGGAATATGTTCAACTGTGAAGTCATCCAAAGGTGTCTCCGGATCAGCGG 180
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
305 TGAGCCCTTATCTTACCTTCAATCCATCTCGGATGTCGCTTGCATTAAGGCCATAA 364
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 TGAGCCCTTATCTTACCTTCAATCCATCTCGGATGTCGCTTGCATTAAGGCCATAA 240
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
365 TGGTTAAAGTGTGGATGAAGCAACCATCTTGAATCTTAATTAAGCAAAACAATGAC 424
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 TGGTTAAAGTGTGGATGAAGCAACCATCTTGAATCTTAATTAAGCAAAACAATGAC 300
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
425 AGCGTCAACAGATCAAGAGCAGCATATCTCAGAGAAACAGAGAAAGCCCTGGATGA 484
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 AGCGTCAACAGATCAAGAGCAGCATATCTCAGAGAAACAGAGAAAGCCCTGGATGA 360
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
485 TGAAGAAAGCCCTTACAGGTCACTTGAAGGATGTTTGTAGTCTGTCTGCTAAAACTCC 544
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 TTAAGAAAGCCCTTACAGGTCACTTGAAGGATGTTTGTAGTCTGTCTGCTAAAACTCC 420
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
545 CGCAATTTGATGCTGATGAATCTGCTGCCATGAGGCGCTTGGAACTGTGAAGATA 604
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 CGCAATTTGATGCTGATGAATCTGCTGCCATGAGGCGCTTGGAACTGTGAAGATA 480
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
605 CTCTAATTGAGATTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTTAAAGGGTCT 664
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 CTCTAATTGAGATTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTTAAAGGGTCT 540
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
665 ACAGAGAGAACTGAAGAGATCTGCCAAAGACATAAATCAGACACATCTGGAGATT 724
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 ACAGAGAGAACTGAAGAGATCTGCCAAAGACATAAATCAGACACATCTGGAGATT 600
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
725 TTCGAAACGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 784
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
601 TTCGAAACGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
785 AAGACTTGGCTGATTCAGATCCAGGCGCTTGTATGAAGCAGAGAAAGAGAAAGGGA 844
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
661 AAGACTTGGCTGATTCAGATCCAGGCGCTTGTATGAAGCAGAGAAAGAGAAAGGGA 720
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
845 CAGACGTAAACGCTTCAATACCATCTTACCACCAAGCTATCCACACTTCGCAGAG 904
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
721 CAGACGTAAACGCTTCAATACCATCTTACCACCAAGCTATCCACACTTCGCAGAG 780
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
905 TGTTCAGAAATACCAAGTACAGTAAAGTACATGAACAAAGTCTTGACCTGGAGT 964
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
781 TGTTCAGAAATACCAAGTACAGTAAAGTACATGAACAAAGTCTTGACCTGGAGT 840
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
965 TGAAGGTGACATTCAGAAATGCTTCAAGTACAGTACAGTACAGTACAGTACAGTACAG 1024
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
841 TGAAGGTGACATTCAGAAATGCTTCAAGTACAGTACAGTACAGTACAGTACAGTACAG 900
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1025 CTTTCTTTGACAGAGCTTCATCAAGCCATGAAAGGTGTTGGAATCTGCCATAGGCAT 1084
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
901 CTTTCTTTGACAGAGCTTCATCAAGCCATGAAAGGTGTTGGAATCTGCCATAGGCAT 960
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1085 TGATCAGATTTATGTTTCCGTTCTGAATTTGATGATGATGATGATGATGATGATGATGAT 1144
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
961 TGATCAGATTTATGTTTCCGTTCTGAATTTGATGATGATGATGATGATGATGATGATGAT 1020
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1145 AGAAGATGATGTTGTTATCTTCCCTTTGCCAAGCCATCTGGATGAAACCAAGAGGAGATG 1204
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1021 AGAAGATGATGTTGTTATCTTCCCTTTGCCAAGCCATCTGGATGAAACCAAGAGGATG 1080
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1205 AGAAATCTCGGTGCTTTTGTGAGGAACTAAACATTCCTTGTGATGTTCTCAAGCTA 1264
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1081 AGAAATCTCGGTGCTTTTGTGAGGAACTAAACATTCCTTGTGATGTTCTCAAGCTA 1140
2y ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1265 TGATCAGAGACTTTTAAATATATTTTCACTCCCTAAGCTTAAATAGGAAAGTTTCTTC 1324
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1141 TGATCAGAGACTTTTAAATATATTTTCACTCCCTAAGCTTAAATAGGAAAGTTTCTTC 1200

```

```

1325 AACAGGATTACAGTGTAGCTACCTACATGCTGAAAAATATAGCCCTTTAAATCATTTTAT 1384
1201 AACAGGATTACAGTGTAGCTACCTACATGCTGAAAAATATAGCCCTTTAAATCATTTTAT 1260
1385 ATTATAACTCTGTATAATAGAGATAAGTCCATTTTAAAAAATGTTTCCCAAAACATA 1444
1261 ATTATAACTCTGTATAATAGAGATAAGTCCATTTTAAAAAATGTTTCCCAAAACATA 1320
1445 AAAACCTATCAAGTGTCTTCTAGTAAACATACATGAGAAAGATGCTATGCTAGCTGAAAA 1504
1321 AAAACCTATCAAGTGTCTTCTAGTAAACATACATGAGAAAGATGCTATGCTAGCTGAAAA 1380
1505 TAAATGNCGTG 1516
1381 TAAATGACGTC 1392

RESULT 3
US-09-641-638-86
; Sequence 86, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Amick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 86
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 10-245-412 : polymorphic base A or G
; NAME/KEY: misc.binding
; LOCATION: 481..500
; OTHER INFORMATION: 10-245-412.mis1, potential
; NAME/KEY: misc.binding
; LOCATION: 502..521
; OTHER INFORMATION: 10-245-412.mis2, potential complement
; NAME/KEY: primer.bind
; LOCATION: 90..107
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer.bind
; LOCATION: 509..528
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc.binding
; LOCATION: 489..513
; OTHER INFORMATION: 10-245-412 potential probe
; NAME/KEY: misc.feature
; LOCATION: 544,552,855..857,865
; OTHER INFORMATION: n=a, g, c or t
;
US-09-641-638-86

Query Match 22.0%; Score 332.8; DB 4; Length 1001;
Best Local Similarity 99.1%; Pred. No. 7.2e-92;
Matches 334; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1180 CTGGATGAAACCAAGGAGAGTATGAGAAATCTCTGGTGGCTCTTTGTGAGGAACTAA 1239  
Db 21 CAGGATGAAACCAAGGAGATATGAGAAATCTCTGGTGGCTCTTTGTGAGGAACTAA 80  
QY 1240 ACATTCCCTTGATGCTCAAGCTATGATCAAGAGACTTTAAATATATATTTTCACTTA 1299  
Db 81 ACATTCCCTTGATGCTCAAGCTATGATCAAGAGACTTTAAATATATATTTTCACTTA 140  
QY 1300 TAAGCTTAAATAGGAAAGTTTCTTCAACAGGATTACAGTGTAGCTACATGCTGAAA 1359  
Db 141 TAAGCTTAAATAGGAAAGTTTCTTCAACAGGATTACAGTGTAGCTACATGCTGAAA 200  
QY 1360 AATATAGCTTTAAATCAATTTTATATATATATATATATATATATATATATATATAT 1419  
Db 201 AATATAGCTTTAAATCAATTTTATATATATATATATATATATATATATATATATAT 260  
QY 1420 TTAATAATGTTTCCCAACCAATTAACCCCTATACAGTGTAGTCTAGTAACTATCATG 1479  
Db 261 TTAATAATGTTTCCCAACCAATTAACCCCTATACAGTGTAGTCTAGTAACTATCATG 320  
QY 1480 AGAAAGATGCTATGATGCTGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1516  
Db 321 AGAAAGATGCTATGATGCTGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 357

## RESULT 4

US-09-023-655-1152  
; Sequence 1152, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Sellhauer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1152:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1362 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: g187137  
; US-09-023-655-1152

Query Match 21.1%; Score 319.4; DB 4; Length 1362;  
Best Local Similarity 59.0%; Pred. No. 1.1e-87;  
Matches 567; Conservative 0; Mismatches 391; Indels 3; Gaps 1;  
QY 302 CGGTGAGCCCTATCTACCTTCAATCCATCTCGATGTGCTGCCCTTGATTAAGGCA 361  
Db 126 CTGTCAAAGCCTATATACTTTGATGCTGAGCGGGATGCTTTGAACATTGAAACAGCA 185  
QY 362 TAATGTTAAAGGTGTGGATGAAGCAACCATCATTGACATTTCTAACTAAGCGAACAATG 421  
Db 186 TCAAGACCAAGGTGTGGATGAGTCAACATTTTGACCAACCGCAGCAATG 245  
QY 422 CACAGGTCACAGATCAAGCAGCATATCTCCAGAAACAGGAAAGCCCTGGATGAAA 481  
Db 246 CACAGAGACAGGATATTCCTTCGCTTACCAGAAAGCAAAAGAACTTGCATCAG 305  
QY 482 CACTGAAGAAAGCCCTTACAGGTCACTTGTAGAGGTTGTTTGTAGCTCTGCTAAAAAATC 541  
Db 306 CACTGAAGTCAAGCTTATCTGCCACCTGGAGACGGTGTATTTTGGGCTTATTGAAGAC 365  
QY 542 CAGCGCAATTTGATGCTGATGAACCTTGTGCTGCTGATGAAGGCTTGGAACTGATGAAG 601  
Db 366 CTGCTCAGTATGACGCTTCTGAGCTAAAGCTTCCATGAAGGGGCTGGGAACCGCAGG 425  
QY 602 ATACTCTAATTCAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGG 661  
Db 426 ACTCTCTCAATGAGATCATCTGCTCCAGAACCAACAGGAGCTGCAGGAAATTAACAGAG 485  
QY 662 TCTACAGAGAGGAACTGAAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAG 721  
Db 486 TCTACAGAGGAAATGTACAGACTGATCTGAGAGAGACATTTATTCGGACACATCTGGTG 545  
QY 722 ATTTTGGGACGCTTGTCTTCTCTGTGTAAGGGTACCGATCTCAGGA--CTTTGGTG 778  
Db 546 ACTTCGCGAAGTGTATGTTGCTGCGCAAGGGTGAAGAGCAGAGATGCTCTGTCA 605  
QY 779 TGAATCAAGACTTTGGCTGATTCAGATGCGCAGGCGCTTGTATGAAGCAGAGAAAGAGAA 838  
Db 606 TTGATTATGAACCTGATTGACCAAGATGCTCGGATCTCTATGACGCTGGAGTGAAGAGA 665  
QY 839 AGGGGACAGACGTAAACGTTTCAATCCATCTTACCACAGAGAGCTATCCACACTTC 898  
Db 666 AAGGAACTGATGTTCCCAAGTGGATCAGCATCATGACCGAGCGGAGCTGCCACCTCC 725  
QY 899 GCAGAGTGTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTTGACC 958  
Db 726 AGAAAGTATTTGATAGGTACAAGGTTACAGCCCTTATGACATGTTGAAAGCATCAGGA 785  
QY 959 TGGAGTTGAAAGGTGACATTAAGAAATGCTCAGACTATCGTGAAGTCGCCCAACAGCA 1018  
Db 786 AAGAGGTTAAAGGAGACCTGGAATGCTTTCTGAACTGTTTCAAGTGTGATTCAGAA 845  
QY 1019 AACGAGCTTTCTTTGAGAGAGAGCTTCAAGCCATGAAGGCTGTTGAACTCGCCATA 1078  
Db 846 AGCCCTGTATTTTGTGCTGATCGGCTGATGACTCCATGAAGGCGAGGAGCGAGATA 905  
QY 1079 AGCATTTGATCAGATTAATGTTTCCGTTCTGAAATTCAGATGAATGATATCAAGCAT 1138  
Db 906 AGGTCTGTATCAGATCATGCTCTCCCGCAGTGAAGTGGACATGTTGAAATTAGTCTG 965  
QY 1139 TCTATCAGAGAGATGATGTTATCTCTCCCTTTGCCAAGCCATCTGAGTGAACCAAGGAG 1198  
Db 966 AATTCAAGAGAAAGTACGCGAAGTCCCTGTACTATTTATCCAGAGACACTAAGGCG 1025  
QY 1199 AGTATGAGAAATCTCTGGTGGCTTTTGTGGAGGAAACTAAACATTCCTTGAATGCTC 1258  
Db 1026 ACTACCAAGAGCGCTGCTGTACTCTGTGTGGAGATGACTGAAGCCGCGACACGCGCTG 1085  
QY 1259 A 1259  
Db 1086 A 1086

## RESULT 5

JS-09-023-655-1149  
 ; Sequence 1149, Application US/09023655  
 ; Patent No. 6607879  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cocks, Benjamin G.  
 ; APPLICANT: Susan G. Stuart  
 ; APPLICANT: Jeffrey J. Sellhamer  
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
 ; TITLE OF INVENTION: EXPRESSION  
 ; NUMBER OF SEQUENCES: 1508  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/023,655  
 ; FILING DATE: HEREWITH  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Zeller, Karen J.  
 ; REGISTRATION NUMBER: 37,071  
 ; REFERENCE/DOCKET NUMBER: PA-0001 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (650) 855-0555  
 ; TELEFAX: (650) 845-4166  
 ; INFORMATION FOR SEQ ID NO: 1149:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1339 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GENBANK  
 ; CLONE: g186967  
 ; JS-09-023-655-1149

Query Match 17.7%; Score 268.4; DB 4; Length 1339;  
 Best Local Similarity 55.2%; Pred. No. 5.2e-72;  
 Matches 524; Conservative 0; Mismatches 426; Indels 0; Gaps 0;

2y	286	GGTGGTCCCGGATCAGCGGTGAGCCCTACCTTCAATCCATCCCTCGGATGCGCT	345
Db	62	GTGGACACCGGAGACAGTAAGATATCCAGACTTAGCCCATCAGTGGATGCTGA	121
2y	346	GCCTTGATAGGCCATAATGGTTAAAGGTGTGATGAAGCAACCATCATTCATCTA	405
Db	122	GCTATTGAGAAAGCAATCAGAGGAATTGGAACCTGATGAGAAATGCTCATCAGCATTCG	181
2y	406	ACTAAGCGMAACATGACAGCGTCAACAGATCAACAGCATATCTCCAGGAAACAGGA	465
Db	182	ACTGAGAGGTCAATGACAGCGGACGCTGATTTTAAAGNAATATCAGCAGCATATGA	241
2y	466	AAGCCCTTGATGAAACACTGAAGAAACCCCTTACAGGTCACTTGAGGAGGTGTTT	525
Db	242	AAGGAGCTGAAAGATGACTTGAAGGGTGAATCTCTCGGCCACTTTGAGCATCTCATGGTG	301
2y	526	GCTGTGCTAAAACATCCAGCGCAATTTGATGCTGATCAATTCGTGCTGCATGAAGGGC	585
Db	302	GCCTTAGTACTCCACCAGCAGTCTTTTATGTCAAAGCAGCTTAAAGAAATCAATGAAGGGC	361

QY	586	CTTGGAACTGATGAAGATACCTTAATTGAGATTTTGGCATCAAGAACTAAACAAGAATC	645
Db	362	CGGGACACAAACGAAGATGCCCTTGATTGAATCTTAATACCAAGCAAGCAGCAATG	421
QY	646	AGAGACATTAAACAGGGTCTACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAACC	705
Db	422	AAGGATATCTCTCAAGCTTATTATACAGTATACAAAGAGAGTCTTGGAGATGACATTAGT	481
QY	706	TCAGACATCTGAGAGATTTTCGGAACGCTTTGCTTTCTTGTAAAGGTGACCGATCT	765
Db	482	TCGGAACATCTGGTGAATCTCCGGAAGCTCTGTTGACTTTTGGCAGATGGCAGAGAT	541
QY	766	GAGGACTTTGGTGTGAATGAAGACTTGGCTGATTTCAGATGCCAGGCTTGTATGAAGCA	825
Db	542	GAAAGTCTGAAGTGGATGAGCATCTGGCCAAACAAAGATGCCAGATCTCTATAAGCT	601
QY	826	GGAGAAAGGAAAGGGGACAGACGTAACGCTGTTCATATACCATCTTACCCAGAGC	885
Db	602	GGTGAGAACAGATGGGGCAGCGATGAAGACAAATTCATCTGAGATCTCTGTTTAAAGGAGC	661
QY	886	TATCCACAACCTTCGACAGAGTGTTCAGAAATACACCAAGTACAGTAAAGCATGACATGAAC	945
Db	662	TTTCTCAATTAATAAATAACATTTGATGATACAGAAATATCACCCAAAGGACATTTGTG	721
QY	946	AAAGTTCTGGACCTGGAGTTGAAAGGTGACATTGAGAAATGCCCTCACAGTATCGTGAAG	1005
Db	722	GACAGCATAAAGGAGAAATTTCTGGGCAATTTTGAAGACTTACTGTTGGCCATAGTTAAT	781
QY	1006	TGCGCCACAAACCAAGCTTTCTTTGCAGAGAGTTTCATCAAGCATCAGAGGTGTT	1065
Db	782	TGTGTGAGGAACACCGCGGCTTTTACCGCAAGACTGCAATCGAGCCTTGAAGGGTATT	841
QY	1066	GGAACTCGCCATAAGGCAATTCATCAGGATTTATGTTTCCCGTCTTGAATTTGACATGAAT	1125
Db	842	CGAAGTGTGATGAGTTTACTCTGAACCGAATAATGTTGCCAGATCAGAAATTCACCTTTG	901
QY	1126	GATATCAAGCATCTATCAGAAAGATGTATGTTATCTCCCTTTCGCAAGCCATCCTGGAT	1185
Db	902	GACATTCGAACAGAGTTCAAGAGCAATTTATGGCTATTCCTTATTCAGCAATTAATTCG	961
QY	1186	GAAACCAAGGAGATGATGAGAAATCTCGTGTGCTCTTTTGTGGAGGAAA	1235
Db	962	GATACTTCTGGAGACTATGAATCACACTCTTAAATCTGTGTGGGAGA	1011

## RESULT 6

US-08-526-136-1  
 ; Sequence 1, Application US/08526136  
 ; Patent No. 6107089  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Towle, Christine A. et al.  
 ; TITLE OF INVENTION: ANNEXIN XI  
 ; NUMBER OF SEQUENCES: 36  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; COMPUTER: IBM PS/2 Model 502 or 55SX  
 ; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
 ; SOFTWARE: WordPerfect (Version 5.0)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/526,136  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/214,036  
 ; FILING DATE:

Db 1408 AAGGCTCACAGGCGCATGAGGGAGCCGGAACCAAGACCGGACCCCTGATCCGCATCA 1467  
 Qy 1097 TGGTTTCCGTTCTGAAATTGACATGAATGATATCAAAAGCATTCTATCAGAAGATGTATG 1156  
 Db 1468 TGGTGTCTCGACCGAGATCGACCTCTCTGGACATCAGACGACAGATATAAGCGGCTGTATG 1527  
 Qy 1157 GTATCTCCCTTTGCCAAGCCATCTCTGGATGAACCAAGGAGAGTATGAGAAAAATCCTGG 1216  
 Db 1528 GCAAGTCGCTGTACCAGCATCACTGAGGAGACACTTCAGGGGATTACCGGAAGATTCTGC 1587  
 Qy 1217 TGGCTCTTTGTGGAGAAA 1235  
 Db 1588 TGAAGATCTGTGGTGCAA 1606  
  
 RESULT 7  
 US-08-526-136-3  
 ; Sequence 3, Application US/08526136  
 ; Patent No. 6107089  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Towle, Christine A. et al.  
 ; TITLE OF INVENTION: ANNEXIN XI  
 ; NUMBER OF SEQUENCES: 36  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson  
 ; CITY: Boston  
 ; STREET: 225 Franklin Street  
 ; STATE: Massachusetts  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; COMPUTER: IBM PS/2 Model 50Z or 55SX  
 ; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
 ; SOFTWARE: WordPerfect (Version 5.0)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/526,136  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/214,036  
 ; FILING DATE:  
 ; APPLICATION NUMBER: 07/837,775  
 ; FILING DATE: February 13, 1992  
 ; APPLICATION NUMBER: 07/764,465  
 ; FILING DATE: September 23, 1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Clark, Paul T.  
 ; REGISTRATION NUMBER: 30,162  
 ; REFERENCE/DOCKET NUMBER: 00786/099001  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 542-5070  
 ; TELEFAX: (617) 542-8906  
 ; TELEX: 200154  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2311  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-526-136-3  
  
 Query Match 16.5%; Score 250.2; DB 3; Length 2311;  
 Best Local Similarity 54.5%; Pred. No. 2.9e-66;  
 Matches 501; Conservative 0; Mismatches 418; Indels 0; Gaps 0  
  
 Qy 317 CTACCTTCAATCCATCTCTCGGATCGCTGCCCTTGCAATAAGGCCATAATGGTTAAAGGTG 376  
 Db 694 CTGGCTTTGACCCCTCGGAGATGCTGAAGTCTGCGGAAGGCCATGAAGGCTTTGGGA 753  
 Qy 377 TGGATGAGCAACCATCATTTGACATTTCTACTAAGCGAAACATGCACAGCGTCAACAGA 436



```

754 CTGACGAGCGGCCATCAATTGACTGCTGGTAGTCGCTCCCAACAGCAGACGAGA 813
437 TCAAGCAGCATATCTCCAGAAACAGAGAAAGCCCTGGATGAAACACTGAGAAAGCCC 496
814 TCCTCTCTGCTTCAAGACAGCATATGGAAGGATTTGATCAAGATCTGAAATCTGAAC 873
497 TTACAGGTCACTTGGAGGCTTTTGTAGTCTGCTTAAATCTCCAGCCCAATTTGATG 556
874 TGTGAGAACTTTGAGAGCAATCTTGCCCTGATGAGACCCCTGTCTCTTTGAGC 933
557 CTGATGAATCTGCTGCGCATGAAGGCTTTGAACTGATGAGATGATCTCTAATTTGAGA 616
934 CTTATGAGATAAGGAAGCTATCAAGGGGGGGGCACTGATGAAGCCTGCTGATCGAGA 993
617 TTTTGGGATCAAGAACTAAACAAAGATCAAGAGCATTTAAGAGGCTCTACAGAGAGAAC 676
994 TCCTGGCTCTCCGAGCAACGAGCAATCCGGAGCTGAACAGATCTCAAGACAGAT 1053
677 TGAAGAGAGATCTGCCCAAGACATACCTCAGACATCTGGAGATTTTCGGAACGCTT 736
1054 TCAAAAGACCCCTGAGGAGGCCATTCGAGCGACACTTCAGGGCACTTCAGCGGCTCC 1113
737 TGTCTCTCTTGAAGGTGACGATCTGAGGACTTTGGTGTGAATGAAGACTTGGCTG 796
1114 TCATCTCTCTCTCAGGAAACCGGATGAAGACAAAGCTGGACATGACCCCTGTGCC 1173
797 ATTGAGATGCCAGGCGCTTGTATGAAGCAGAGAGAGAGAAAGGGGACAGACGTAACG 856
1174 AGAGAGATGTCAGGAGCTCTATGAGCTGGGAGAACCGCTGGGAAACAGATGAGTCCA 1233
857 TGTTCATATACCATCTTACCAAGAGAGCTATCCACACTTCGAGAGTGTTCAGAAAT 916
1234 AGTTCATGCGATCTGTGCTCCCGAGCGGGCCACCTGGTGGCAGTTTAAACGAGT 1293
917 ACACCAAGTACAGTAAGCATGACATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGTGACA 976
1294 ATCAGAGGATGACAGGAGCTGACATTTGAGAGAGCATCTGCGGGAGATGTCCGGGAGC 1353
977 TTGAGAAATGCTCAGACTATCGTGAAGTGGGCAACAGAAACACGCTTCTTTTCAG 1036
1354 TGGAGCAGGCTGCTGCTGTGGTGAATGCTTTAAGATACCCAGCCTTCTTTCTG 1413
1037 AGAAGCTTCATCAAGCCATGAAGGTGTGGAATCTCGCCATGAAGGCTATGATCAGGATTA 1096
1414 AAGGCTCAACAGGCTATGAGGAGCGGAAACCAAGACCGGACCTGATCGCATCA 1473
1097 TGGTTCCGCTTCTGAATGACATGAATGATATCAAGGATTTCTATCAGAGATGATG 1156
1474 TGGTGTCTCGCAGCAGATCGACCTCTCTGACATCAGAGCAGAGATATAAGCGGCTGTATG 1533
1157 GPATCTCCTTTGCCAAGCCTCTCTGATGAAACCAAGAGGAGATGATGAGAAATCTCTGG 1216
1534 GCAAGTCTGTATCACGACATCACGGAGACATTCAGGGGATTCAGGAGATTTCTGC 1593
1217 TGGCTCTTTTGGAGGAA 1235
1594 TGAAGATCTGTGGTGGCAA 1612

```

```

RESULT 8
US-09-023-655-1190
; Sequence 1190, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE

```

```

CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1190:
SEQUENCE CHARACTERISTICS:
LENGTH: 1301 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 5189616
US-09-023-655-1190

Query Match 16.2%; Score 245.2; DB 4; Length 1301;
Best Local Similarity 54.3%; Pred. No. 6.8e-65;
Matches 496; Conservative 0; Mismatches 418; Indels 0; Gaps 0;

QY 322 TTCATCCATCTCCGATGTCCTGCTTGCATAGGCCATTAATGTTAAAGGTGCGAT 381
DB 119 TTCAATGCCATGAGATGCCAGACCTGAGGAGGCCNTGAAAGGCTCGGCCCGAT 178
QY 382 GAAGCAACCATCATTTGACATTTCTAAAGCGAAACAATGACAGCGTCAACAGATCAAA 441
DB 179 GAAGACGCAATTATTAGCGTCTTGCTACCGCAACACCGCCAGCGCAGGAGATCAGG 238
QY 442 GCAGCATATCTCCAGGAACAGAAAGCCCTGGATGAAACACTGAAGAAAGCCCTTACA 501
DB 239 ACAGCCTACAAGAGCACCATCGCAGGAGCTTGTAGCGACCTGAAGTCAGACTGAGT 298
QY 502 GGTCACTTGGAGAGGCTGTTTTAGCTCTGCTTAAAGAACTCCAGCGCAATTTGATGCTGAT 561
DB 299 GGCACACTTCGAGCAGGTGTTTGGGGATGATGAGCCCAACCGTCTGATGACGTGCAA 358
QY 562 GAACTTGTGCTGCCATGAAGGCTTGGAACTGTGATGAGATGATCTAATTCAGATTTTG 621
DB 359 GAGCTGCAAGGGGCCATGAAGGAGCGCGCACTGATGAGGCTGCTTAATGAGATCTCTG 418
QY 622 GCATCAAGAACTAAACAAAGAAATCAGAGACATTTAAGAGGCTCTACAGAGAGAACTGAAG 681
DB 419 GCCTCCGGNCCCTCAGGAGATCGCGCCATAGCCAAACCTACCAGCAGCAATATGA 478
QY 682 AGAGATCTGCCCAAGACATACCTCAGACATCTGGAGATTTTCGGAACGCTTTGCTT 741
DB 479 CGGAGCCTTGAAGATGACATTCGCTCTGACATCGTTTCCATGTTCCAGGAGTGTGGTG 538
QY 742 TCTCTTGTGAAGGCTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTGATTCA 801
DB 539 TCTCTGTGAGTGTGGGAGGATGAAGAAATTAATCTGAGCATGCTCTCGTGAGACAG 598
QY 802 GATGCCAGGCGCTTGTATGAAGCAGAGAAAGAGAAAGGAGACAGACGTTAAACGTTGTT 861

```

```

599 GATGCCAGGACCTGTATGAGCTGGAGAGAAATGGGGACAGATGAGGTGAATTT 658
862 AATACATCCTTACCAACAGAGCTATCCAACTTCGAGAGTGTTCAGAAATACACC 921
659 CTAACCTGTTCTCTGTTCCCGAACCGAAATACCTGTTGCCATGTTGATGAATACAAA 718
922 AAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGGTGACATTGAG 981
719 AGGATATACAGAGGATATGAAAGAGATTAATCTGAACATCTGATAGCTTGA 778
982 AATGCTCAGAGCTATCGTGAAGTGGCGCAACAGCAACCAAGCTTTCTTGCAGAGAG 1041
779 GATGCTCTCTGCTGCTATAGTAAAGTGCATGAGAAACAAATCTGCATATTTTCTGCAAAAG 838
1042 CTTCTACAGCCATGAAGGTGTGAACTCCGCAATAGGCAATTCATCAGGATATGTT 1101
839 CTCATAAATCGATGAAGGCTTGGCCACCGATATACACCTCATCAGATGATGTT 898
1102 TCCGCTTCTGAATGACATGAATGATATCAAAAGCATTTCTACAGAGATGATGTTATC 1161
899 TCTCGAGCAGAAATGACATGTTGGATATCCGGGCACACTTCAAGAGACTCTATGGAAG 958
1162 TCCCTTTGCCAGCCATCTCGTGAAGTGAACCAAGAGAGATGAGAAATCTGTTGCT 1221
959 TCTCTGATCTCGTTCAAGGCTGACATCTGAGACTACAGGAAGTACTGCTTGT 1018
1222 CTTTGTGGAGAAA 1235
1019 CTCTGGAGGAGA 1032

```

RESULT 9

```

US-08-125-746-4
; Sequence 4, Application US/08125746
; Patent No. 5591633
; GENERAL INFORMATION:
; APPLICANT: SAINO, YUSHI
; APPLICANT: IWASAKI, AKIO
; APPLICANT: SUDA, MAKOTO
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,746
; FILING DATE: 24-SEP-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/807,623
; FILING DATE: 13-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 037227/1987
; FILING DATE: 20-FEB-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 184428/1987
; FILING DATE: 23-JUL-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5591633man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-074-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000

```

```

; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 957 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-125-746-4

Query Match 15.8%; Score 239.8; DB 1; Length 957;
Best Local Similarity 53.3%; Pred. NO. 2.5e-63;
Matches 505; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

289 GGTCCCGGATCAGCGGTGAGCCCTATCTACCTTCAATCCATCTCTCGGATGTCGTGCC 348
7 GTTCTCAGAGGCACTGACTGACTTCCCTGGATTGATGAGCGGCTCATGCGAAGT 66
349 TTGCATAAGGCCATATGTTTAAAGTGTGGATGAAGCAACCATCATTCATTCACATCTTA 408
67 CTTCCGGAAGCTATGAAGGCTTTGGGCACAGATGAGGAGCATCTCTGACTCTGTGTGACA 126
409 AAGCGAAACAATCCACAGCGTCAACAGATCAAAAGCAGCATATCTCCAGGAAACAGGAAAG 468
127 TCCGGAAGTATGCTCAGCGCCAGGAAATCTCTGCACTTTTAAAGACTCTGTTTGCAGG 186
469 CCCCCTGGATGAACACTGAAGAAAGCCCTTACAGTCACTCTGAGGAGGTGTTTGTAGCT 528
187 GATCTTCTGATGACCTGAAATCAGAACTAACTGGAATAATTTGAAAAATTTAATTTGGCT 246
529 CTGCTAAAACCTCCAGCGCAATTTGATGCTGATGAATCTCGTCTGCCATGAAGGCGCTT 588
247 CTGATGAACCCCTCTCGGCTTTATGATGCTTATGAACTGAACATGCTTTGAAGCGAGCT 306
589 GGAACCTGATGAAGATACTCTAATTGAGATTTTGGCATCAAGAACTAACAAGAAATCAGA 648
307 GGAACAAATGAAAAGTACTGACAGAAATTTATGCTTCAAGGACACCTGGAAGAACTGAGA 366
549 GACATTAACAGGCTCTACAGAGAGGAACTGAAGAGATCTGGCCAAAGACATTAACCTCA 708
367 GCATCAACAAGTTTATGAAGAAGATATGGCTCAAGCTGAAGATGACGTGTTGGGG 426
709 GACACATCTGGAGATTTTCGAAACGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 768
427 GACACTTCAGGCTACTACAGAGCGATTTGGTGGTTCTCTCTTCTTCTTCTTCTTCTTCT 486
769 GACTTTGGTGTGAATGAGACTTTGGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGA 828
487 GATGCTGGAATGATGAAGCTCAAGTTGAACAAGATGCTCAGGCTTTTATTTTCAAGCTTGA 546
829 GAAAGGAGAAAGGGGACAGACGTAACGCTGTTTCAATACCAATCTTACCACCAAGCTAT 888
547 GACTTAAATGGGGACAGATGAAGAAAGTTTATACCACTTTTGGAAACACCAAGTGTG 606
889 CCACAACTTCGAGAGTGTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAA 948
607 TCTCATTTGAGAAAGGTGTTTGAACAAGTACATGACTATATCAGGATTTCAATTTAGGAA 666
949 GTTCTGGACCTGGAGTTGAAAGGTGACATTAAGAAATGCTTCAAGCTATCGTGAAGTGC 1008
667 ACCATTGACCGGAGACTTCTGGCAATTTAGAGCACTACTCTCTGCTGTTGTGAAATCT 726
1009 GCCAAGCAAGCAACAGCTTTCTTTGAGAGAGCTTCAATCAAGCCATGAAGGTGTTGGA 1068
727 ATTCGAAGTATACCTGCTACCTTTCAGAGAGACCTCTTATTTATGTAAGGAGGCTGGG 786
1069 ACTCGCATTAAGGCATTTGATCAGGATTTATGTTTCCGTTTCCGTTTCTGAAATGACATGA 1128
787 ACAGATGATCATACCTTCATCAGAGTCACTGTTTCCAGGAGTGAATGATGCTGTTTAA 846
1129 ATCAAGCATTTCTATCAAGAGATGATATGTTATCTCCCTTTGCCAAGCCATCTCGATGA 1188

```

Db 847 ATCAGGAAGGAGTTAGGAAGAAATTTTCCACCTCTCTTTATTCATGATTAAGGGAGAT 906  
Qy 1189 ACCAAAGGAGAGTATGAGAAATCTCTGTGCTCTTTGTGGAGAAA 1235  
Db 907 ACATCTGGGACTATAGAAAGCTCTTCTGCTGCTCTGTGGAGA 953

RESULT 10  
5225537-3  
; Patent No. 5225537  
; APPLICANT: FOSTER, DONALD  
; TITLE OF INVENTION: METHODS FOR PRODUCING HYBRID  
; PHOSPHOLIPID-BINDING PROTEINS  
; NUMBER OF SEQUENCES: 14  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/459,082  
; FILING DATE: 29-DEC-1989  
; SEQ ID NO:3:  
; LENGTH: 1460  
5225537-3

Query Match 15.8%; Score 239.8; DB 6; Length 1460;  
Best Local Similarity 53.3%; Pred. No. 3.3e-63;  
Matches 505; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

Qy 289 GGTCCCGATCAGCGGTGAGCCCTATCTACCTTCAATCCATCCTCGGATGCTGCC 348  
Db 22 GTTCTCAGAGGACGTGACTGACTTCTCCGTGATTTGATGAGCGGTGATGAGAACT 81  
Qy 349 TTGCATAAGGCATATATGTTAAAGGTGTGATGAGCAACCATCATGATTAAC 408  
Db 82 CTTCCGAAGGCTATGAAAGGCTTGGGCACAGATGAGGAGCATCTGACTCTCTTGACA 141  
Qy 409 AAGCGAACAATGACAGGCTCAGAGATCAAGAGCATATCTCCAGGAACAGGAAG 468  
Db 142 TCCGAAGTAACTGTCAGCGCCAGGAATCTCTGAGCTTTTAAGACTCTGTTGGCAGG 201  
Qy 469 CCCCTGGATGAACACTGAAGAAACCCCTTACAGTCACTTGGAGAGGTTGTTAGCT 528  
Db 202 GATCTTCTGGATGACCTGAATCAGAACTAACTGGAAATTTGAAATTAATTTGGCT 261  
Qy 529 CTGCTAAACCTCAGCGCAATTTGATCTGATGACTTCTGCTGCTGATGAGGCTT 588  
Db 262 CTGATGAACCTCTCGGCTTTATGCTTATGATGACTTGAACATGCTGCTGAGGAGCT 321  
Qy 589 GGAATGATGAAGATCTTAATGAGATTTTGGCATCAAGAACTTAAACAAAGAAATCAGA 648  
Db 322 GGAACAAATGAAGATGACTGACAGAAATTTGCTTCAAGGACACCTGAAGAACTGAGA 381  
Qy 649 GACATTAACAGGCTTACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATACCTCA 708  
Db 382 GCCATCAACAAAGTTTATGAAGAAATATGCTCAAGCCTTGAAGATGACGCTGCTGGG 441  
Qy 709 GACATCTGGAGATTTTCGGAACTTTGCTTCTTCTTGAAGGTCACCGATCTGAG 768  
Db 442 GACATTCAGGCTACTACAGCGGATGTTGGTGTCTCTCTCAGCTTAACAGAGCCCT 501  
Qy 769 GACTTGGTGTGAATGAAGACTTGGCTGATTCAGATGCAAGGCTTGTATGAAGCAGGA 828  
Db 502 GATGCTGGAATGATGAAGCTCAAGTTGAACAAGATGCTCAGGCTTATTTTCAGGCTGA 561  
Qy 829 GAAAGGAGAAAGGACAGCACTTAACTGTTCAATACCTTACCTTACCACAGAGCTAT 888  
Db 562 GAATTAATGGGGACAGATGAAGAAAGTTTATCAACATCTTTGGAACAGAAAGTGTG 621  
Qy 889 CCACAACCTTCGAGAGTGTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAA 948  
Db 622 TCTCATTTTGAAGAAAGTGTTCACAAGTACATGACTATATCAGGATTTCAAATGAGGAA 681  
Qy 949 GTTCTGGACTGGATGAAAGGTGACATTTGAGAAATGCTTACAGCTATCGTGAAGTGC 1008  
Db 682 ACCATTGACCGGAGACTTCTGGCAATTTAGAGCAACTACTCTTGTGTTGAAATCT 741

Qy 1009 GCCACAAGCAACACAGCTTTCTTTCAGAGAGCTTCTATCAAGCCATGAAAGGCTTGA 1068  
Db 742 ATTCAAGTATACCTGCTACCTTTCAGAGACCTCTTATATGCTTATCAAGGAGCTGG 801  
Qy 1069 ACTCGCCATAAGGCATTGATCAGGATTTATGGTTTCCCGTCTTGAATTCACATGAATGAT 1128  
Db 802 ACAGATGATCATACCTCATCAGACTCATGGTTCCAGAGTGAGATTGATCTGTTTAA 861  
Qy 1129 ATCAAAGCAATTCATCAGAGATGATGATCTCCCTTTCAGAGCCATCTCTGATGAA 1188  
Db 862 ATCAGGAAGGAGTTTAGGAAGAAATTTGCCACCTCTCTTTATTCATGATTAAGGAGAT 921  
Qy 1189 ACCAAAGGAGATGATGAGAAATCTCTGTTGGCTCTTTGTGGAGAAA 1235  
Db 922 ACATCTGGGACTATAGAAAGCTCTTCTGCTGCTCTGTGGAGAAGA 968

RESULT 11  
US-08-125-746-2  
; Sequence 2, Application US/08125746  
; Patent No. 5591633  
; GENERAL INFORMATION:  
; APPLICANT: SAINO, YUSHI  
; APPLICANT: INASAKI, AKIO  
; APPLICANT: SUDA, MAKOTO  
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDE  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/125,746  
; FILING DATE: 24-SEP-1993  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/807,623  
; FILING DATE: 13-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 037227/1987  
; FILING DATE: 20-FEB-1987  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 184428/1987  
; FILING DATE: 23-JUL-1987  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5591633man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-074-0 DIV  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1567 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 136...1095  
US-08-125-746-2

```

Query Match      15.8%; Score 239.8; DB 1; Length 1567;
Best Local Similarity 53.3%; Pred. No. 3.5e-63;
Matches 505; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

QY 289 GGTCCCGGATCAGCGGTGAGCCCTATCTACCTTCAATCCATCTCCGATGTCGTGCC 348
DB 145 GTTCTCAGAGGCACTGTGACTGCTCCCTGATTTGATGAGCGGCTGATGCGAAACT 204
QY 349 TTGCATAGGCCATATGTTAAAGGTGGATGAAGCAACCATCATTTGACATTTCACT 408
DB 205 CTTTGGAGGCTATGAAGGCTTGGCAGATGAGGAGCATCTGACTCTGTGACA 264
QY 409 AAGCGAAACAATGCAAGCGTCAACAGATCAAGAGCAGATATCTCCAGAAACAGAAAG 468
DB 265 TCCGGAAGTAATGCTCAGCGCCAGGAAATCTCTGAGCTTTTAAGACTCTGTTGGCAGG 324
QY 469 CCCCTGGATGAACACTGAAGAACCCCTTACAGTCACTTGGAGGCTGTTTACCT 528
DB 325 GATCTTCTGGATGACCTGAAATCAGAACTAACTGGAATAATTGAAATAATTATGCTGCT 384
QY 529 CTGCTAAAACCTCCAGCGCAATTTGATGCTGATGAACCTTCGTGCTGCCATGAAGGCCCTT 588
DB 385 CTGATGAACCCCTCTCGGCTTTATGATGCTTATGAATGAACATGAACATGCTTGAAGGGAGCT 444
QY 589 GGAACGTGATGATGACTCTTAATGATGATTTGGCATCAAGAACTAAACAAGAAATCAGA 648
DB 445 GGAACAAATGAAAGATGCTGACAGAAATTAATGCTTCAAGGACACCTGAAGAACTGAGA 504
QY 649 GACATTAACAGGGTCTACAGAGAGGAACTGAAGAGAGATCTGCGCCAAAGACATACCTCA 708
DB 505 GCATCAACAAGTTTATGAAGAAGATATGGCTCAAGCTGGAAGATGACGTGGTGGG 564
QY 709 GACACATCTGGAGATTTTGGAAAGCTTTGCTTTCTTCTTGAAGGTCACCGATCTGAG 768
DB 565 GACACTTACGGGTACTACAGCGGATGTTGGTGGTTCTCTTCCAGGCTAACAGAACCTT 624
QY 769 GACTTGGTGTGAATGAAGACTTGGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGA 828
DB 625 GATGCTGGAATGATGAAGCTCAAGTTGAACAGATGCTCAGGCTTTATTCAGGCTGA 684
QY 829 GAAAGAGAAAGGGGACAGACGTAAAGCTGTTTCAATACCATCTTACACAGAGAGCTAT 888
DB 685 GAACTTTAAATGGGGGACAGATGAAGAAAGTTTATCACCCTCTTTGGAACAGAAAGTGT 744
QY 889 CCACAACTTCGACAGTGTGTTTCAAGAAATACACCAAGTACAGTAAGCATGACATGAACAAA 948
DB 745 TCTCATTTGGAAGGTTGTTGACAGTACATGACTATATCAGGATTTCAAATTCAGGAA 804
QY 949 GTTCTGACCTGGATGGAAGGTGACATGAGAAATGCTCAGACTATCTGTAAGTGC 1008
DB 805 ACCATTGACCGGAGACTTCTGGCAATTTAGAGCAACTACTCTCTGCTGTTGAAATCT 864
QY 1009 GCCAAGCAACCCAGCTTTCTTTGAGAGAGCTTTCATCAAGCCATGAAGGTGTTGA 1068
DB 865 ATTCGAAGTATACCTGCTACCTTTGAGAGACCTCTATATGCTATGAGGAGCTGG 924
QY 1069 ACTCGCCATAGGCAATGATCAGATTAATGTTTCCGTTCTGAAATTAACATGATAT 1128
DB 925 ACAGATCATACCTCATCAGAGTCAATGTTTCCAGGAGTGAATGATCTGTTTAAAC 984
QY 1129 ATCAAGCATCTATCAGAGATGATGTTATCTCCCTTTGCCAGCCATCTCGATGAA 1188
DB 985 ATCAGGAAGGAGTTAGGAAGAATTTTGGCACTCTCTTTATTCATGATTAAGGAGAT 1044
QY 1189 ACCAAAGGAGATGATGAGAAATCTCGTGGCTCTTTTGTGGAGAAA 1235
DB 1045 ACATCTGGGACTATAAGAAAGCTCTTCTGCTGCTCTGTTGAGAGA 1091

```

RESULT 12  
 US-09-324-096A-1  
 ; Sequence 1, Application US/09324096A  
 ; Patent No. 6323313

```

; GENERAL INFORMATION:
; APPLICANT: Tait, Jonathan
; TITLE OF INVENTION: ANNEXIN DERIVATIVE WITH ENDOGENOUS CHELATION SITES
; FILE REFERENCE: US/09-324-096A
; CURRENT APPLICATION NUMBER: US/09-324-096A
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(981)
; US-09-324-096A-1

Query Match      15.6%; Score 236.6; DB 4; Length 981;
Best Local Similarity 53.1%; Pred. No. 2.5e-62;
Matches 503; Conservative 0; Mismatches 444; Indels 0; Gaps 0;

QY 289 GGTCCCGGATCAGCGGTGAGCCCTATCCCTCAATCCATCTCCGATGTCGTGCC 348
DB 31 GTTCTCAGAGGCACTGTGACTGACTTCCCTGGATTTGATGAGCGGCTGATGCGAAACT 90
QY 349 TTGCATAGGCCATATGTTAAAGGTGGATGAAGCAACCATCATTTGACATTTCAACT 408
DB 91 CTTTGGAGGCTTATGAAGGCTTGGGACAGATGAGGAGCATCTGACTCTGTGACA 150
QY 409 AAGCGAAACAATGCAAGCGTCAACAGATCAAAAGCAGCATATCTCCAGGAAACAGAAAG 468
DB 151 TCCGGAAGTAAATGCTCAGCGCCAGGAAATCTCTGACGCTTTTAAGACTCTGTGTCAGG 210
QY 469 CCCCTGATGAACACTGAAGAAAGCCCTTACAGGTCACCTTCAGGAGGTGTTTGTAGCT 528
DB 211 GATCTTCTGGATGACCTGAAATCAGAACTAACTTGGAAATTTGAAATAATTATTTGGCT 270
QY 529 CTGTAAAAAATCCAGCGCAATTTGATGCTGATGAACCTTCGCTGCCATGAAGGCGCTT 588
DB 271 CTGATGAACCCCTCTCGGCTTTATGATGCTTATGAACTAAGAAACATGCTTTGAAGGAGCT 330
QY 589 GGAACATGATGAGTACTCTTAATGAGATTTTGGCATCAAGAACTAAACAAGAAATCAGA 648
DB 331 GGAACAAATGAAAGAGTACTGACAGAAATTTATGCTTCAAGGACACCTGAGAACTGAGA 390
QY 649 GACATTAACAGGGTCTACAGAGAGGAACTGAAGAGAGATCTGGCCAAAGACATAAACCCTCA 708
DB 391 GCCATCAACAAGTTTATGAAGAAGATATGGCTCAAGCTGGAAGATGACGTGGTGGG 450
QY 709 GACACATCTGGAGATTTTGGAAAGCTTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 768
DB 451 GACACTTCAGGCTACTACAGCGGATGTTGGTGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 510
QY 769 GACTTTGGTGTGAATGAAGACTTGGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGA 828
DB 511 GATCTGGAATGATGAGCTCAAGTTGAACAGATGCTCAGGCTTTATTTTCAAGCTGGA 570
QY 829 GAAAGGAGAAAGGGAACAGAGCTAAACGTTGTTCAATACCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 888
DB 571 GAACTTAAATGGGGGACAGATGAAGAAAGTTTATCACCCTTTTGGAAACAGAAAGTGT 630
QY 889 CCACAACTTCGACAGGTGTTTCAGAAATACACCAAGTACAGTAAGCATGATGAACGAA 948
DB 631 TCTCATTTGAGAAAGGTTTGAACAGTACATGACTATATCAGGATTTCAAATTCAGGAA 690
QY 949 GTTCTGGACCTGGAGTTGAAAGGTGACATTTGAGAAATGCTTCAAGATTCAGCTATGCTGAAAGTGC 1008
DB 691 ACCATTGACCGGAGACTTCTGGCAATTTTAGAGCAACTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 750
QY 1009 GCCCAAGCAACAGCTTTCTTTGAGAGAGAGCTTCTCATCAAGCCATGAAGAGGTGTTGGA 1068
DB 751 ATTCGAAGTATACCTGCTACCTTTCAGAGAGAGCCCTCTTATATGCTATGAGGAGAGCTGGG 810

```

3603 CTATATGCAATTAACCTGTTTGTGTAACCTCGATCTTCCTTTCCGATACCTATTGAT 3662  
 128867 CTATATGCAATTAACCTGTTTGTGTAACCTCGATCTTCCTTTCCGATACCTATTGAT 128808  
 3663 GAATAAAGAAATTAAGTGA 3682  
 128807 GAATAAAGAAATTAAGTGA 128788

RESULT 5  
 AC068844 170233 bp DNA linear HTG 24-AUG-2002  
 LOCUS Homo sapiens chromosome 3, clone RP11-690F8 map 3, WORKING DRAFT  
 DEFINITION SEQUENCE, 23 unordered pieces.  
 AC068844 2 GI:8516083  
 VERSION HTG; HTGS PHASE1; HTGS\_DRAFT.  
 KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (Bases 1 to 170233)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bada, F.,  
 Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,  
 Campiolo, A., Casle, A., Choepel, Y., Colangelo, M., Collins, S.,  
 Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand-pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
 Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
 Melidrim, J., Menus, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J.,  
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
 O'Neill, D., Oliver, J.M., Oliver, J., Peterson, K., Pierse, N.,  
 Piani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talanas, J.,  
 Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

2 (Bases 1 to 170233)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bada, F.,  
 Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,  
 Campiolo, A., Casle, A., Choepel, Y., Colangelo, M., Collins, S.,  
 Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand-pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
 Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
 Melidrim, J., Menus, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J.,  
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
 O'Neill, D., Oliver, J.M., Oliver, J., Peterson, K., Pierse, N.,  
 Piani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talanas, J.,  
 Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission  
 Submitted (10-MAY-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (Bases 1 to 170233)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bada, F.,  
 Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,  
 Campiolo, A., Casle, A., Choepel, Y., Colangelo, M., Collins, S.,  
 Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand-pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
 Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
 Melidrim, J., Menus, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J.,  
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
 O'Neill, D., Oliver, J.M., Oliver, J., Peterson, K., Pierse, N.,  
 Piani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talanas, J.,  
 Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE

JOURNAL  
 COMMENT

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jun 14, 2000 this sequence version replaced gi:7767838.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7518

Center clone name: 690 F 8

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156993 bases at least Q40

Consensus quality: 163905 bases at least Q30

Consensus quality: 166725 bases at least Q20

Insert size: 24400; agarose-fp

Insert size: 168033; sum-of-contigs

Quality coverage: 3.8 in Q20 bases; agarose-fp

Quality coverage: 5.6 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 23 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1957: contig of 1957 bp in length  
 \* 1958 2057: gap of 100 bp  
 \* 2058 3880: contig of 1823 bp in length  
 \* 3881 3980: gap of 100 bp  
 \* 3981 5189: contig of 1209 bp in length  
 \* 5190 5289: gap of 100 bp  
 \* 5290 7410: contig of 2121 bp in length  
 \* 7411 7510: gap of 100 bp  
 \* 7511 8754: contig of 1244 bp in length  
 \* 8755 8954: gap of 100 bp  
 \* 8955 10300: contig of 2046 bp in length  
 \* 10301 11000: gap of 100 bp  
 \* 11001 14582: contig of 3682 bp in length  
 \* 14583 14782: gap of 100 bp  
 \* 14783 19532: contig of 4750 bp in length  
 \* 19533 19632: gap of 100 bp  
 \* 19633 25241: contig of 5609 bp in length  
 \* 25242 25341: gap of 100 bp  
 \* 25342 32065: contig of 6724 bp in length  
 \* 32066 32165: gap of 100 bp  
 \* 32166 40193: contig of 8028 bp in length  
 \* 40194 40293: gap of 100 bp  
 \* 40294 46461: contig of 6168 bp in length  
 \* 46462 46561: gap of 100 bp  
 \* 46562 54554: contig of 8093 bp in length  
 \* 54555 54754: gap of 100 bp  
 \* 54755 63205: contig of 8451 bp in length  
 \* 63206 69683: contig of 6378 bp in length  
 \* 69684 69783: gap of 100 bp  
 \* 69784 77173: contig of 7390 bp in length  
 \* 77174 77273: gap of 100 bp  
 \* 77274 86740: contig of 9467 bp in length  
 \* 86741 86840: gap of 100 bp  
 \* 86841 94331: contig of 7491 bp in length  
 \* 94332 94431: gap of 100 bp  
 \* 94432 101503: contig of 7072 bp in length  
 \* 101504 101603: gap of 100 bp  
 \* 101604 118443: contig of 16840 bp in length

\* 118444 118543: gap of 100 bp  
 \* 118544 132371: contig of 13828 bp in length  
 \* 132372 132471: gap of 100 bp  
 \* 132472 153434: contig of 20963 bp in length  
 \* 153435 153535: gap of 100 bp  
 \* 153535 170233: contig of 16699 bp in length.

# FEATURES

Source  
 1. 170233  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /map="3"  
 /clone="RP11-690F8"  
 /clone\_1lb="RPCT-11 Human Male BAC"  
 1. 1957  
 /note="assembly\_fragment"  
 2058. 3880  
 /note="assembly\_fragment"  
 3981. 5189  
 /note="assembly\_fragment"  
 clone\_end="77"  
 vector\_side="right"  
 5290. 7410  
 /note="assembly\_fragment"  
 7511. 8754  
 /note="assembly\_fragment"  
 8855. 10900  
 /note="assembly\_fragment"  
 11001. 14682  
 /note="assembly\_fragment"  
 14783. 19532  
 /note="assembly\_fragment"  
 19633. 25241  
 /note="assembly\_fragment"  
 clone\_end="86"  
 vector\_side="right"  
 25342. 32065  
 /note="assembly\_fragment"  
 32166. 40193  
 /note="assembly\_fragment"  
 40294. 46461  
 /note="assembly\_fragment"  
 46562. 54654  
 /note="assembly\_fragment"  
 54755. 63205  
 /note="assembly\_fragment"  
 63306. 69683  
 /note="assembly\_fragment"  
 69784. 77173  
 /note="assembly\_fragment"  
 77274. 86740  
 /note="assembly\_fragment"  
 86841. 94331  
 /note="assembly\_fragment"  
 94432. 101503  
 /note="assembly\_fragment"  
 101604. 118443  
 /note="assembly\_fragment"  
 118544. 132371  
 /note="assembly\_fragment"  
 132472. 153434  
 /note="assembly\_fragment"  
 153535. 170233  
 /note="assembly\_fragment"

Query Match 73.4%; Score 2710.4; DB 2; Length 170233;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2714; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 963 ACGAAACAGCTGAAAGACCAAGAGAGTGAAGCCAAAGAGATCTTAATGAGAGGCAAG 1022  
 Db 149284 ACGAAACAGCTGAAAGACCAAGAGAGTGAAGCCAAAGATCTTAATGAGAGGCAAG 149343

QY	1023	ACCAAGTCTGTGATCCAGCATCTTCTCACTTGAAGCGCTGAGTTCAATACAACTG	1082
Db	149344	ACCAAGTCTGTGATCCAGCATCTTCTCACTTGAAGCGCTGAGTTCAATACAACTG	149403
QY	1083	TCGTGACAGTGGCCAAATCAACAGTATTGCTTAATATAGCAATGAGTTTATTTGTTT	1142
Db	149404	TGTGTACAGTGGCCAAATCAACAGTATTGCTTAATATAGCAATGAGTTTATTTGTTT	149463
QY	1143	ATTGTTTTCGAATTAAGATTTGAAGTGGCTGCTTGAAGGAGGCAACGCTT	1202
Db	149464	ATTGTTTTCGAATTAAGATTTGAAGTGGCTGCTTGAAGGAGGCAACGCTT	149523
QY	1203	CATTTCTGAGAGTCTTTAAGAAATCTGAATATGCTCTGAGGCTGAGGCTAGTAA	1262
Db	149524	CATTTCTGAGAGTCTTTAAGAAATCTGAATATGCTCTGAGGCTGAGGCTAGTAA	149583
QY	1263	GGAAACTGCATCAGATTTGAAAGAGAAACAGCCAAATCTGAACCTCTTTGAGTTAC	1322
Db	149584	GGAAACTGCATCAGATTTGAAAGAGAAACAGCCAAATCTGAACCTCTTTGAGTTAC	149643
QY	1323	TGCATCTGTGACAGAGCTGACAGGAGTGACACATGACAGAGAACTTAGACAGGCTGT	1382
Db	149644	TGCATCTGTGACAGAGCTGACAGGAGTGACACATGACAGAGAACTTAGACAGGCTGT	149703
QY	1383	CCCCGAGAGAGGTTTGGGAAAGCTCCACGAGAGAAAGCTCTGCTTCCAGCCTTT	1442
Db	149704	CCCCGAGAGAGGTTTGGGAAAGCTCCACGAGAGAAAGCTCTGCTTCCAGCCTTT	149763
QY	1443	TGCATTCGGCTGAGATACAGACCTCCAGCATCCAGGCATCTTGTGCTCCCAATACG	1502
Db	149764	TGCATTCGGCTGAGATACAGACCTCCAGCATCCAGGCATCTTGTGCTCCCAATACG	149823
QY	1503	CCTCTAGATACATAGCCATCTCTAGTTAACCAAGTCTCCTGACCTGGATGAGTT	1562
Db	149824	CCTCTAGATACATAGCCATCTCTAGTTAACCAAGTCTCCTGACCTGGATGAGTT	149883
QY	1563	TCTGGAGAGTACACCCAAATGATACATCTTGTATATTTGAGCCCTTAGAGGACT	1622
Db	149884	TCTGGAGAGTACACCCAAATGATACATCTTGTATATTTGAGCCCTTAGAGGACT	149943
QY	1623	AACCAATTTTAAAAATCTTTTTCACAAAGTGTATTTCTGTAAACACTTTT	1682
Db	149944	AACCAATTTTAAAAATCTTTTTCACAAAGTGTATTTCTGTAAACACTTTT	150003
QY	1683	TGCAAGTGTACTTATCTTCAATATTAATTAATTAATTTTAAATATTTAT	1742
Db	150004	TGCAAGTGTACTTATCTTCAATATTAATTAATTAATTTTAAATATTTAT	150063
QY	1743	TTTCTTGAAGTATTAAGCTTTTGTATTAATTTTGAAGTATCCACCACTTATAG	1802
Db	150064	TTTCTTGAAGTATTAAGCTTTTGTATTAATTTTGAAGTATCCACCACTTATAG	150123
QY	1803	TGGAAGAGTTTGGGTTTCTTCTGTGAGGAGGCTGAATTAACCAATGCTCCACCC	1862
Db	150124	TGGAAGAGTTTGGGTTTCTTCTGTGAGGAGGCTGAATTAACCAATGCTCCACCC	150183
QY	1863	TGCAATCTAGATGACAGCCCATGATGAGCCCTTACGCTTCCAGAGTCCACTATG	1922
Db	150184	TGCAATCTAGATGACAGCCCATGATGAGCCCTTACGCTTCCAGAGTCCACTATG	150243
QY	1923	CCAGAGAGCAAGGTCCTTGAAGCTTAAGCCAGGAGAAAGACATCTTCAATAAAACT	1982
Db	150244	CCAGAGAGCAAGGTCCTTGAAGCTTAAGCCAGGAGAAAGACATCTTCAATAAAACT	150303
QY	1983	TTCAAGATCAAACTTAATTTTATTTATTTCTGAGAGTGAAGCAATCAGTAT	2042
Db	150304	TTCAAGATCAAACTTAATTTTATTTATTTCTGAGAGTGAAGCAATCAGTAT	150363
QY	2043	TCCCAAGATGGGACAGAGGCAACCAAGAGGCTTGAAGATCTCCAGCCTTACCAATAT	2102
Db	150364	TCCCAAGATGGGACAGAGGCAACCAAGAGGCTTGAAGATCTCCAGCCTTACCAATAT	150423



2y	2103	GTCTATTGCACTAACTAGAGGGTGAGTTGCCCTGTCTCTTCTTTTCTGGAGCTCAG	2162
Db	150424	GCTCATTTCGACTAACTAGAGGGTGAGTTGCCCTGTCTCTTCTTTTCTGGAGCTCAG	150483
2y	2163	TTTCTCTCAGTGAGCTGGTTAAGAATGCATCAACCTTTTGAATTCATGAAGTTATAAAATCTTG	2222
Db	150484	TTTCTCTCAGTGAGCTGGTTAAGAATGCATCAACCTTTTGAATTCATGAAGTTATAAAATCTTG	150543
2y	2223	TGGTTCTGTATCATTTGGTCCAGAGGGAGATAGGTTCCGTGTGAATTTTCCCTTCTTCTCTAT	2282
Db	150544	TGGTTCTGTATCATTTGGTCCAGAGGGAGATAGGTTCCGTGTGAATTTTCCCTTCTTCTCTAT	150603
2y	2283	AGNATAAAATGCAATCTTGTTACTAGAACACAGAAATGTCAGATGCCCAAAAACAGATGAC	2342
Db	150604	AGNATAAAATGCAATCTTGTTACTAGAACACAGAAATGTCAGATGCCCAAAAACAGATGAC	150663
2y	2343	CAGATTTGATCTCAGCCTGATGACCTACAGGTCGTCTATGATATGGAGTCCTCATGGG	2402
Db	150664	CAGATTTGATCTCAGCCTGATGACCTACAGGTCGTCTATGATATGGAGTCCTCATGGG	150723
2y	2403	TAAAGCAGAGAGAGATGGGAAGAGACCAACCCACTCTGTCTTCATATTTTGCATTTC	2462
Db	150724	TAAAGCAGAGAGAGATGGGAAGAGACCAACCCACTCTGTCTTCATATTTTGCATTTC	150783
2y	2463	TGTTTAACTCCGGCTGGAAATAGAAAAACATTCCTCTAGAGATGAGGATAAAAAAGT	2522
Db	150784	TGTTTAACTCCGGCTGGAAATAGAAAAACATTCCTCTAGAGATGAGGATAAAAAAGT	150843
2y	2523	TCAGATTCAAAGGGGGAGAAAAATGGAGATTTAATCTCTAAACTGTGACTTGGGGAGT	2582
Db	150844	TCAGATTCAAAGGGGGAGAAAAATGGAGATTTAATCTCTAAACTGTGACTTGGGGAGT	150903
2y	2583	CAGTCATTTACAGTTAGTCTGTGTCTTTCCAGCTTCTGTGATTAATAACCCCACTCAC	2642
Db	150904	CAGTCATTTACAGTTAGTCTGTGTCTTTCCAGCTTCTGTGATTAATAACCCCACTCAC	150963
2y	2643	CCCTGTTTCAGATGCAATTTGGAAATACCAAGATTTAAATCCCTTGACATPAAGATCTCAT	2702
Db	150964	CCCTGTTTCAGATGCAATTTGGAAATACCAAGATTTAAATCCCTTGACATPAAGATCTCAT	151023
2y	2703	CAGAAAGCAGATTAAGACCAATCAGAGGAAATTTATTTAGTTGTAATGCACAGGCAACT	2762
Db	151024	CAGAAAGCAGATTAAGACCAATCAGAGGAAATTTATTTAGTTGTAATGCACAGGCAACT	151083
Qy	2763	GTGAGAAACTGTGTGCCAAAATAGAAATCCCTCTAGTTTTCTGTCTCTCATTTGAAA	2822
Db	151084	GTGAGAAACTGTGTGCCAAAATAGAAATCCCTCTAGTTTTCTGTCTCTCATTTGAAA	151143
Qy	2823	GGAGAAAAATCCACTTTGTTTAGCAATTTCAAGCTTTTATGTAATCCATCCCATCTAA	2882
Db	151144	GGAGAAAAATCCACTTTGTTTAGCAATTTCAAGCTTTTATGTAATCCATCCCATCTAA	151203
Qy	2883	TCCTTCAAACTCCACTTGTTCAGTCTGAAATGCAGTCCCTGTCCAAGTGCTTGGAGAAC	2942
Db	151204	TCCTTCAAACTCCACTTGTTCAGTCTGAAATGCAGTCCCTGTCCAAGTGCTTGGAGAAC	151263
Qy	2943	TCACAGCAGACGCCCTTAATCAAAAGTTTATCCAGCCCTTGGACATATGGGAGAGGGC	3002
Db	151264	TCACAGCAGACGCCCTTAATCAAAAGTTTATCCAGCCCTTGGACATATGGGAGAGGGC	151323
Qy	3003	AAGAGTACACCAATTTGTTTAAAGCAAGAAACCAAGTGCTCTTCTACTAGTCAATTTAG	3062
Db	151324	AAGAGTACACCAATTTGTTTAAAGCAAGAAACCAAGTGCTCTTCTACTAGTCAATTTAG	151383
Qy	3063	ACATGTTATTCATCCAAGACTACTCTACCTGCAACATTTGAATCTCCAAGAGCAAAATCA	3122
Db	151384	ACATGTTATTCATCCAAGACTACTCTACCTGCAACATTTGAATCTCCAAGAGCAAAATCA	151443
Qy	3123	CATTCTCTTGAGTCTGAGCTCTGTGTGAATATAGGCAGCTGTCGTCTATGCCCTAG	3182
Db	151444	CATTCTCTTGAGTCTGAGCTCTGTGTGAATATAGGCAGCTGTCGTCTATGCCCTAG	151503
Qy	3183	ATCACATGATCTGAGGACCAATTCATGGAAGCTGTGTAATATAGCCTTAGTCTGGGAGTCTTC	3242

Db	151504	ATCACATGATCTGAGGACCAATTCATGGAAGCTGCTAAATAGCCCTAGTCTCGGGAGTCTTC	151563
Qy	3243	CATAAAGTTTTGCATGTGAGCAAAACAAACAGGATTAAACTAGGTTTGGTTCTCTCAGCCCT	3302
Db	151564	CATAAAGTTTTGCATGTGAGCAAAACAAACAGGATTAAACTAGGTTTGGTTCTCTCAGCCCT	151623
Qy	3303	CTAAAGCATGAGGCTTAGCCTGACAGGCTTCCTTGGGCTTCTCTGTGTGTGTAGTTTGG	3362
Db	151624	CTAAAGCATGAGGCTTAGCCTGACAGGCTTCCTTGGGCTTCTCTGTGTGTGTAGTTTGG	151683
Qy	3363	TAAACACTATAGCATCTGTTTAAAGATCCAGTGTCCATGGAACCTTCCACATGCGTGAC	3422
Db	151684	TAAACACTATAGCATCTGTTTAAAGATCCAGTGTCCATGGAACCTTCCACATGCGTGAC	151743
Qy	3423	TCTGACTATATCATGTTTTTGGGAAGCAGGCTTCTCTGCTGTCTTAAACAGCCACGTGG	3482
Db	151744	TCTGACTATATCATGTTTTTGGGAAGCAGGCTTCTCTGCTGTCTTAAACAGCCACGTGG	151803
Qy	3483	ACCACTCTGAATGCTTTTCCCTTTACACCTATGTTTTTAAAGTAGTCAAACTTCAAGAAACA	3542
Db	151804	ACCACTCTGAATGCTTTTCCCTTTACACCTATGTTTTTAAAGTAGTCAAACTTCAAGAAACA	151863
Qy	3543	ATCTAAACAAGTTTCTGTGTGCATATGTTTTGGAACCTTGTATTTAGTATTTAGTAGGCTT	3602
Db	151864	ATCTAAACAAGTTTCTGTGTGCATATGTTTTGGAACCTTGTATTTAGTATTTAGTAGGCTT	151923
Qy	3603	CTATATTGCATTAACTTGTGTTTTTGTAACTCCTGATCTTCTCTTTCGGATACATTGAT	3662
Db	151924	CTATATTGCATTAACTTGTGTTTTTGTAACTCCTGATCTTCTCTTTCGGATACATTGAT	151983
Qy	3663	GAATAAAGAAATTAAGTGA	3682
Db	151984	GAATAAAGAAATTAAGTGA	152003
RESULT 6			
AF288537			
LOCUS		2180 bp	mRNA linear PRI 04-FEB-2001
DEFINITION		Homo sapiens follistatin-related protein mRNA, partial cds.	
ACCESSION		AF288537	
VERSION		AF288537.1	GI:12658308
KEYWORDS			
SOURCE		Homo sapiens (human)	
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
SOURCE			
CDS			

## ORIGIN

Query Match 56.6%; Score 2089; DB 9; Length 2180;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 2163; Conservative 0; Mismatches 10; Indels 8; Gaps 6;

QY 628 AACGATCCAGACCGAGAGAACAAAGTGTGAGGGGACTCTGTGTGATGCTCAT 587  
 DB 1 AACGATCCAGACCGAGAGAACAAAGTGTGAGGGGACTCTGTGTGATGCTCAT 60  
 QY 688 TGAAGTGTGATGAGAAATGCTGATGGAATCAAGCTTCCAGAGTTTCTCAAGTCC 747  
 DB 61 TGAAGTGTGATGAGAAATGCTGATGGAATCAAGCTTCCAGAGTTTCTCAAGTCC 120  
 QY 748 CAACCATTTTTCMAACCTCTCTGAGAGAGATGTGCTCGAGATTAAGCTATGCA 807  
 DB 121 CAACCATTTTTCMAACCTCTCTGAGAGAGATGTGCTCGAGATTAAGCTATGCA 180  
 QY 808 TGAAGTGTGAGACCGAGAGTGAAGTGAACCGCTGTGTGCTGTGAGAAATGAGTCTG 867  
 DB 181 TGAAGTGTGAGACCGAGAGTGAAGTGAACCGCTGTGTGCTGTGAGAAATGAGTCTG 240  
 QY 868 TACAGCCATGACCTGTGACGAGAAAGATCAGAAAGGAGCCAGACCCAGACAGAGAG 927  
 DB 241 TACAGCCATGACCTGTGACGAGAAAGATCAGAAAGGAGCCAGACCCAGACAGAGAG 300  
 QY 928 GATGACCCAGATATGTCCAGAGAGCTCCAAAAGCATAGAGAAACAGTGAACCAAGAG 987  
 DB 301 GATGACCCAGATATGTCCAGAGAGCTCCAAAAGCATAGAGAAACAGTGAACCAAGAG 360  
 QY 988 AGTGAACACCAAGAGATTAATGAAGAGCAACAGACAGTGTCTGATGCCAGATCTT 1047  
 DB 361 AGTGAACACCAAGAGATTAATGAAGAGCAACAGACAGTGTCTGATGCCAGATCTT 420  
 QY 1048 CTCACCTTAGGCGCTGAGTTCAGTATACAAAGTGTGCTACAGTGGCCAAATACAG 1107  
 DB 421 CTCACCTTAGGCGCTGAGTTCAGTATACAAAGTGTGCTACAGTGGCCAAATACAG 480  
 QY 1108 TATTTGCTTATATAGCAATGAGTTTATTTTGTGTTTTCATTAAGATATGA 1167  
 DB 481 TATTTGCTTATATAGCAATGAGTTTATTTTGTGTTTTCATTAAGATATGA 540  
 QY 1168 AGTGTGCTGTGAGAGAGAGAGGCGCAAGCTTCATTTCTAGAGAGCTTTAGAGAA 1227  
 DB 541 AGTGTGCTGTGAGAGAGAGAGGCGCAAGCTTCATTTCTAGAGAGCTTTAGAGAA 600  
 QY 1228 ACTGTAAATGTGTCTGTGGGCTGAGAGCTAGTAAGAACTGCATCAAGATTGAAG 1287  
 DB 601 ACTGTAAATGTGTCTGTGGGCTGAGAGCTAGTAAGAACTGCATCAAGATTGAAG 660  
 QY 1288 GAAACAGACCCAAATCTGAACTCTTTTGAATTACGTCTGTCAAGAGCTGAGAGGA 1347  
 DB 661 GAAACAGACCCAAATCTGAACTCTTTTGAATTACGTCTGTCAAGAGCTGAGAGGA 720  
 QY 1348 GTGACACAGATGCGAGAGAGAACTTGAAGAGGTGTCCCGAGAGAGAGGTTTGGAGCT 1407  
 DB 721 GTGACACAGATGCGAGAGAGAACTTGAAGAGGTGTCCCGAGAGAGAGGTTTGGAGCT 780  
 QY 1408 CCAAGAGAGAGAAAGCTCTCTGTCTTCCAGCTCTTTTCCATTGCGCTGACAGTGAAG 1467  
 DB 781 CCAAGAGAGAGAAAGCTCTCTGTCTTCCAGCTCTTTTCCATTGCGCTGACAGTGAAG 840  
 QY 1468 TCCAGATCCAGCATCTCTGTGTCCCAATTAATGCTCTAGATACATAGCATCTCTCT 1527  
 DB 841 TCCAGATCCAGCATCTCTGTGTCCCAATTAATGCTCTAGATACATAGCATCTCTCT 900  
 QY 1528 AGTTAAACCAAGTGTCCCTGAGACTTGAAGATGATTTCTGAGAGGATACACCAAAATGATG 1587  
 DB 901 AGTTAAACCAAGTGTCCCTGAGACTTGAAGATGATTTCTGAGAGGATACACCAAAATGATG 960  
 QY 1588 CAGATCTGTATACCTTGAAGGCTTGAAGGACCTTAACCAAAATTTTAAATTAATCTTTTA 1647  
 DB 961 CAGATCTGTATACCTTGAAGGCTTGAAGGACCTTAACCAAAATTTTAAATTAATCTTTTA 1020

QY 1648 CCAAGAGTGTCTATTTCTGTGAAACACTTTTGTGGCAAGTGACTTATCTTCAT 1707  
 DB 1021 CCAAGAGTGTCTATTTCTGTGAAACAC-TTGTGTGGCAAGTGACTTATCTTCAT 1079  
 QY 1708 TATTAATCATTAATTAATGTTTATTAATTTATTTCTGACAGGATTAAGCTTTT 1767  
 DB 1080 TATTAATCATTAATTAATGTTTATTAATTTATTTCTGACAGGATTAAGCTTTT 1139  
 QY 1768 GTATTAATTTTTCAGTATGTCACCACTTCATAGTGAAGAGATTTGGGGTCTTCTCG 1827  
 DB 1140 GTATTAATTTTTCAGTATGTCACCACTTCATAGTGAAGAGATTTGGGGTCTTCTCG 1199  
 QY 1828 GTGAGGGGCTGAATTAACCCAGATGTCGCCACCTGCGACATATGATGAGGCCATA 1887  
 DB 1200 GTGAGGGGCTGAATTAACCCAGATGTCGCCACCTGCGACATATGATGAGGCCATA 1259  
 QY 1888 GTTGGCCCCCTAGCTTCCAGCATGTCATATCTGACAGAGAGCAAGGTGCTTAAAC 1947  
 DB 1260 GTTGGCCCCCTAGCTTCCAGCATGTCATATCTGACAGAGAGCAAGGTGCTTAAAC 1319  
 QY 1948 CTAGCCAGAGGAGAGAGATC-TTCATTAATAACCTTCAAGATCCAAATTAATTTGT 2006  
 DB 1320 CTAGCCAGAGGAGAGAGATCTTTCATTAATAACCTTCAAGATCCAAATTAATTTGT 1379  
 QY 2007 TTTTATTTATTTGAGAGATGAGCAATCCAGATATCCCAAGATGCGCAAGAGCCAG 2066  
 DB 1380 TTTTATTTATTTGAGAGATGAGCAATCCAGATATCCCAAGATGCGCAAGAGCCAG 1439  
 QY 2067 CCAAGAGGGCTT-AGGATATCCAGCCCTACCAATATGCTCAATGAGTCACT-AGGAG 2124  
 DB 1440 CCAAGAGGGCTTAAATTTCCAGCCCTACCAATATGCTCAATGAGTCACT-AGGAG 1499  
 QY 2125 GTGAGTGGCCCTGTCTCTCTTTTCT-AGACCTCACTTCTCA---GTGAGCTGT 2180  
 DB 1500 GTGAGTGGCCCTGTCTCTCTTTTCTTTTGGAGCCCTCACTTCTCAAGTGAAGCTGT 1559  
 QY 2181 AAGATGCACTAACCTTTGATTTGATTAAGTATTAATCTGTGCTGATCATTTGTC 2240  
 DB 1560 AAGATGCACTAACCTTTGATTTGATTAAGTATTAATCTGTGCTGATCATTTGTC 1619  
 QY 2241 CAGAGGGAGATAGTGTCTGTGATTTTCTCTCTCTCTATAGATTAATGAATCTTG 2300  
 DB 1620 CAGAGGGAGATAGTGTCTGTGATTTTCTCTCTCTCTCTATAGATTAATGAATCTTG 1679  
 QY 2301 TTAATGACCAAGAAATGTCAGATGCGCCAAAACAAATGACCAATTTGATCTACGCT 2360  
 DB 1680 TTAATGACCAAGAAATGTCAGATGCGCCAAAACAAATGACCAATTTGATCTACGCT 1739  
 QY 2361 GATGACCTTACAGTGTGCTATGATATGAGAGTCTCATGGTAAAGCAGAGAGAGTG 2420  
 DB 1740 GATGACCTTACAGTGTGCTATGATATGAGAGTCTCATGGTAAAGCAGAGAGAGTG 1799  
 QY 2421 GAAAAGAGAACCAACCCCACTCTGTCTTCAATATTTGATTTCACTTCCGCTGG 2480  
 DB 1800 GAAAAGAGAACCAACCCCACTCTGTCTTCAATATTTGATTTCACTTCCGCTGG 1859  
 QY 2481 AAATAGAAAGCATTTCCCTTGAATGAGATTAAGAAAGTTTCAAGTTCAACAGGGGGA 2540  
 DB 1860 AAATAGAAAGCATTTCCCTTGAATGAGATTAAGAAAGTTTCAAGTTCAACAGGGGGA 1919  
 QY 2541 AGAAATGAGATTTAATCTTAATACTGTGACTTGGGAGAGTCAATCATTTTACAGTTAGT 2600  
 DB 1920 AGAAATGAGATTTAATCTTAATACTGTGACTTGGGAGAGTCAATCATTTTACAGTTAGT 1979  
 QY 2601 CCTGTGCTTTCAGCTTCTGTGATTTAATACCCCACTCACTACCTGTTTCAAGTCAAT 2660  
 DB 1980 CCTGTGCTTTCAGCTTCTGTGATTTAATACCCCACTCACTACCTGTTTCAAGTCAAT 2039  
 QY 2661 TGAATACCAAGATTAATCTTGAATGAATGATTCATTTGAGAGAGAGATTAAGA 2720  
 DB 2040 TGAATACCAAGATTAATCTTGAATGAATGATTCATTTGAGAGAGAGATTAAGA 2099



Db 1641 TTAATAATACATTTTACCAAGGTGCTATTCTCTGTGAAAACAC-TTTTTTGGCAAGTT 1699  
QY 1692 GACTTTATTTCTCAATTAATATCATTAATATATGTTTTTTAATATTTATTTCTTGAC 1751  
Db 1700 GACTTTATTTCTCAATTAATATCATTAATATATGTTTTTTAATATTTATTTCTTGAC 1759  
QY 1752 TAGGTATTAAGCTTTTGAATATATTTTTCAGTAGTCCACACATTCATAGTGAAGGAG 1811  
Db 1760 TAGGTATTAAGCTTTTGAATATATTTTTCAGTAGTCCACACATTCATAGTGAAGGAG 1819  
QY 1812 TTTGGGGTTCTTCGNGTGCAGGGGCTGAATAACCCAGATGTCACACCTGCCACATA 1871  
Db 1820 TTTGGGGTTCTTCGNGTGCAGGGGCTGAATAACCCAGATGTCACACCTGCCACATA 1879  
QY 1872 CTAGATGAGCCCATAGTGTGGCCCCCTTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG 1931  
Db 1880 CTAGATGAGCCCATAGTGTGGCCCCCTTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG 1939  
QY 1932 CAAGGGTCCCTTAGACCTTAAGCCAGGGGAAGAAGCATCTT 1971  
Db 1940 CAAGGGTCCCTTAGACCTTAAGCCAGGGGAAGAAGCATCTT 1979

RESULT 8  
HSU06863 1987 bp mRNA linear PRI 11-MAY-1995  
LOCUS Human follistatin-related protein precursor mRNA, complete cds.  
DEFINITION  
ACCESSION U06863  
VERSION U06863.1 GI:536897  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1987)  
Zwijssen,A., Blochx,H., Van Arnhem,W., Willems,J., Franssen,L.,  
Devos,K., Raymakers,J., Van de Voorde,A. and Slegers,H.  
Characterization of a rat C6 glioma-secreted follistatin-related  
protein (FRP). Cloning and sequence of the human homologue  
Eur. J. Biochem. 225 (3), 937-946 (1994)  
JOURNAL 95045570  
MEDLINE 7957230  
PUBMED  
REFERENCE 2 (bases 1 to 1987)  
AUTHORS Zwijssen,A.  
TITLE Direct Submission  
SUBMITTED (17-FEB-1994) An Zwijssen, Biochemistry, University of  
Antwerp UTA, Universiteit Antwerpen, Wilrijk, Antwerp, B-2610, Belgium  
JOURNAL Location/Qualifiers  
FEATURES  
source  
1..1987  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_lines="Hs683"  
/cell\_type="glioma"  
/tissue\_type="brain"  
1..91  
5'UTR  
CDS  
92..1018  
/codon\_start=1  
/product="follistatin-related protein precursor"  
/protein\_id="AAA66062.1"  
/db\_xref="GI:536898"  
/translation="MWKELALALVAVWVRAEELRSKSKICANVFCAGRECAV  
TEKSEPTCLCEQCKPXPVCGSNGKTYLNHCLHRDACLTSKIQVDYDGHCKEKK  
SVSPASPVVQVQNSDELRRITQWLEAIIIPGWFSKGSNYSLEIDKFKQNDNG  
SRLDSEFLKVFQNETAIRITTYDQENKLLRGLCVDAIILIEDNADWKLSPQF  
LKLNPSNPPEKKCALEDEYDVCNRCVACGNGVWVCTAMTCGKQKQKGAQ  
TQTEEMTRYVQELQKHQETAETKRVSTKEI"  
92..151  
sig\_peptide  
152..1015  
mat\_peptide  
/product="follistatin-related protein"  
1019..1987  
3'UTR  
ORIGIN

Query Match 52.5%; Score 1940; DB 9; Length 1987;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1954; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
QY 12 CTCGCCACCTCCGCTTACAGCTCGCTGCCCGCGTCCCTGCCCGCGCCGCCAGGAGACCTGG 71  
Db 21 CTCGCCACCTCCGCTTACAGCTCGCTGCCCGCGTCCCTGCCCGCGCCGCCAGGAGACCTGG 80  
QY 72 ACCAGACCAACGATGTGGAAACGCTCGCTCGCGCTCGCGCTCGCGCTGGTGGCGGTGCGCT 131  
Db 81 ACCAGACCAACGATGTGGAAACGCTCGCTCGCGCTCGCGCTCGCGCTGGTGGCGGTGCGCT 140  
QY 132 GGGTCCGCGCCGAGAGAGCTAAGAGCAATCAAGATCTGTGCCAATGTGTTTGTG 191  
Db 141 GGGTCCGCGCCGAGAGAGCTAAGAGCAATCAAGATCTGTGCCAATGTGTTTGTG 200  
QY 192 GAGCCGCGCGGGAATGTGCAGTCAAGAGAGGAGGAAACCCACCTGTCTCTGCATTGAGC 251  
Db 201 GAGCCGCGCGGGAATGTGCAGTCAAGAGAGGAGGAAACCCACCTGTCTCTGCATTGAGC 260  
QY 252 AATGCAACCTCACAAGAGGCTGTGTGGCAGTAATGGCAGACCTACCTCAACCACT 311  
Db 261 AATGCAACCTCACAAGAGGCTGTGTGGCAGTAATGGCAGACCTACCTCAACCACT 320  
QY 312 GTGAATGTCATCGAGATGCTGCTCCTCACTGGATCCAAATCCAGGTTGATTACGATGGAC 371  
Db 321 GTGAATGTCATCGAGATGCTGCTCCTCACTGGATCCAAATCCAGGTTGATTACGATGGAC 380  
QY 372 ACTGCAAGAGAGAAATCCGTAAGTCCATCTGCCAGCCAGTGTGTTGCTATCAGTCA 431  
Db 381 ACTGCAAGAGAGAAATCCGTAAGTCCATCTGCCAGCCAGTGTGTTGCTATCAGTCA 440  
QY 432 ACCGTGATGAGTCCGAGCTGCATCATCTCAGTGCCTGGAGCTCAGATCATCTCAGATG 491  
Db 441 ACCGTGATGAGTCCGAGCTGCATCATCTCAGTGCCTGGAGCTCAGATCATCTCAGATG 500  
QY 492 GCTGGTCTCTAAAGCAGCAACTACAGTGAATCCTAGACAGATTTTAAAGACTTTG 551  
Db 501 GCTGGTCTCTAAAGCAGCAACTACAGTGAATCCTAGACAGATTTTAAAGACTTTG 560  
QY 552 ATAAATGGTATTTCTCGCTGGACTCCAGTGAATCTCAGTGAATCTTGAAGTTTGTGAAACAGATGAAA 611  
Db 561 ATAAATGGTATTTCTCGCTGGACTCCAGTGAATCTCAGTGAATCTTGAAGTTTGTGAAACAGATGAAA 620  
QY 612 CTGCCATCAATATTACAGTATCCAGACCCAGAGAACCAACAGTTGCTTAGGGAGCTCT 671  
Db 621 CTGCCATCAATATTACAGTATCCAGACCCAGAGAACCAACAGTTGCTTAGGGAGCTCT 680  
QY 672 GTGTTGATGCTCTCATTTGAACCTGTCTGATGAAAATGCTGATGAAAATCTCAGCTTCCAG 731  
Db 681 GTGTTGATGCTCTCATTTGAACCTGTCTGATGAAAATGCTGATGAAAATCTCAGCTTCCAG 740  
QY 732 AGTTTCTCAAGTGGCTCAACCCATCTTCAACCCCTCCTGAGAAAGTGTGCCCTGGAGG 791  
Db 741 AGTTTCTCAAGTGGCTCAACCCATCTTCAACCCCTCCTGAGAAAGTGTGCCCTGGAGG 800  
QY 792 ATGAACAGTATCCAGATGGAGCTGAGACCGAGGTGGACTGTAAACCGCTGTGTCTGTGCT 851  
Db 801 ATGAACAGTATCCAGATGGAGCTGAGACCGAGGTGGACTGTAAACCGCTGTGTCTGTGCT 860  
QY 852 GTGGAATTTGGTCTGTGACAGCCATGACCTGTGACGGAAGAAATCAGAAAGGGGGGCCAGA 911  
Db 861 GTGGAATTTGGTCTGTGACAGCCATGACCTGTGACGGAAGAAATCAGAAAGGGGGGCCAGA 920  
QY 912 CCAGAGCAGAGAGAGATGACAGATATGTCCAGAGCTCAAAAGATCAGGAAACAG 971  
Db 921 CCAGAGCAGAGAGAGATGACAGATATGTCCAGAGCTCAAAAGATCAGGAAACAG 980  
QY 972 CTGAAAAGACCAAGAGAGTGAGCAACCAAGAGATCTAATGAGAGGCGCAGACCAAGTGT 1031  
Db 981 CTGAAAAGACCAAGAGAGTGAGCAACCAAGAGATCTAATGAGAGGCGCAGACCAAGTGT 1040

[illegible]

**ORGANISM** Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

**REFERENCE**  
**AUTHORS** Garcia,T., roman Roman,S., Baron,R., Cail,K., Theilhaber,J.,  
Connolly,T., Jackson,A., Bushnell,S.E. and Rawadi,G.  
**TITLE** Genes involved in osteogenesis, and methods of use  
**JOURNAL** Patent: WO 02081745-A 77 17-OCT-2002,  
Aventis Pharma S.A. (FR),  
Location/Qualifiers  
source 1. .2538  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

**FEATURES**

**ORIGIN**

Query Match	46.8%; Score 1729.8; DB 6; Length 2538;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 1742;	Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY      12 CTCGCCACCTCCGTTACAGTGTGCCTCGCGCGCTCTGCCCGCGCGCCCCCAGAGACCTGG 71
       |||
Db       6 CTCGCCACCTCCGTTACAGTGTGCCTCGCGCGCTCTGCCCGCGCGCCCCCAGAGACCTGG 65
       |||

QY      72 ACCAGACACCATGTGGAACGCTGCCTCGCGCTCGCGCTGC CGCTGTGTCGCGTGCCT 131
       |||
Db       66 ACCAGACACCATGTGGAACGCTGCCTCGCGCTCGCGCTGC CGCTGTGTCGCGTGCCT 125
       |||

QY      132 GGATCCCGCGCGGAGGAGACTTAAGGACAATCCAAGANTCTGTGCCCAATCTGTTTTGTG 191
       |||
Db       126 GGATCCCGCGCGGAGGAGACTTAAGGACAATCCAAGANTCTGTGCCCAATCTGTTTTGTG 185
       |||

QY      192 GAGCGCGCGCGGAATGTGCAGTCACAGAAAGGGGAACCCACCCTGTCTCTGCATTGAGC 251
       |||
Db       186 GAGCGCGCGCGGAATGTGCAGTCACAGAAAGGGGAACCCACCCTGTCTCTGCATTGAGC 245
       |||

QY      252 AATCGAAACCTCACAGAGCGCTGTGTGGCAGTAATGGCGAGACCTACCTCAACCACT 311
       |||
Db       246 AATCGAAACCTCACAGAGCGCTGTGTGGCAGTAATGGCGAGACCTACCTCAACCACT 305
       |||

QY      312 GTGAACCTGCATCGAGATGCTCGCTCACCTGGATCCAAAATCCAGGTTCATTCAGATGAC 371
       |||
Db       306 GTGAACCTGCATCGAGATGCTCGCTCACCTGGATCCAAAATCCAGGTTCATTCAGATGAC 365
       |||

QY      372 ACTCGAAAGAAGAAATCGTAAGTCCATCTGCCAGCGCCAGTTGTGTGCTATCAGTGCA 431
       |||
Db       366 ACTCGAAAGAAGAAATCGTAAGTCCATCTGCCAGCGCCAGTTGTGTGCTATCAGTGCA 425
       |||

QY      432 ACCGTGATGAGCTCCGACGTCGCATCATCCAGTGTGGAAGCTGAGACTCAGATCATCTCAGATG 491
       |||
Db       426 ACCGTGATGAGCTCCGACGTCGCATCATCCAGTGTGGAAGCTGAGACTCAGATCATCTCAGATG 485
       |||

QY      492 GCTGTTCTCTTAAAGCGAGCAACTACAGTGAATACTTAGCAAAGTATTTTAAAGAACCTTTG 551
       |||
Db       486 GCTGTTCTCTTAAAGCGAGCAACTACAGTGAATACTTAGCAAAGTATTTTAAAGAACCTTTG 545
       |||

QY      552 ATAATGTTGATTTCTCGCTGGACTCCAGTGAATTCCTGAACTTTGTGGAACAGATGA 611
       |||
Db       546 ATAATGTTGATTTCTCGCTGGACTCCAGTGAATTCCTGAACTTTGTGGAACAGATGA 605
       |||

QY      612 CTGCCATCAATATTACAACTATCCAGACCCAGGAGAACAAACAAAGTTGCTTAGGGGACTCT 671
       |||
Db       606 CTGCCATCAATATTACAACTATCCAGACCCAGGAGAACAAACAAAGTTGCTTAGGGGACTCT 665
       |||

QY      672 GTGTTGATGCTCTCATTTGAACCTGTCTGATGAAAAATGCTGATTGGAACCTCAGCTCCNAG 731
       |||
Db       666 GTGTTGATGCTCTCATTTGAACCTGTCTGATGAAAAATGCTGATTGGAACCTCAGCTCCNAG 725
       |||

QY      732 AGTTTCTCAAGTGCTCTCAACCCATCTTTCAAACCTCTCTGAGAAAGGTGTGCGCTTGAGG 791
       |||
Db       726 AGTTTCTCAAGTGCTCTCAACCCATCTTTCAAACCTCTCTGAGAAAGGTGTGCGCTTGAGG 785
       |||

QY      792 ATGAAAAGTATGCAATGAGAGCTGAGACCGAGGTGGAATGTAACCGCTGTGCTGTGCT 851
       |||

```





b	246		AATGCAAACTCACAAGAGGCGCTGTGTGTGGCAGTAAATGCGAAGACCTACTCTCAACCACT	305
y	312		GTGAACTGCATCGAGATGCTGCCTCACTGGATCCAAAATCCAAAGTTGATTACGATGGAC	371
b	306		GTGAACTGCATCGAGATGCTGCCTCACTGGATCCAAAATCCAAAGTTGATTACGATGGAC	365
y	372		ACTGAAAAGAGAAGAAATCGTAAAGTCCATCTCTGCCAGCCGAGTTGTTGCTATCACTCCA	431
b	366		ACTGAAAAGAGAAGAAATCGTAAAGTCCATCTCTGCCAGCCGAGTTGTTGCTATCACTCCA	425
y	432		ACCGTGATGAGCTCCGACGTGCGATCATCAGTGGGTGGAAGCTGAGATCATCTCAGATG	491
b	426		ACCGTGATGAGCTCCGACGTGCGATCATCAGTGGGTGGAAGCTGAGATCATCTCAGATG	485
y	492		GCTGGTTCTCTAAAGGCGAGCAATACAGTGAATCCTAGACAAGTATTTTAAAGACTTTG	551
b	486		GCTGGTTCTCTAAAGGCGAGCAATACAGTGAATCCTAGACAAGTATTTTAAAGACTTTG	545
y	552		ATAATGGTGATTCTCGCTCGAATCCAGTGAATTCCTGAAGTTTGTGGAAACAGAAATGAA	611
b	546		ATAATGGTGATTCTCGCTCGAATCCAGTGAATTCCTGAAGTTTGTGGAAACAGAAATGAA	605
y	612		CTGCCATCAATATTACAAAGTATCCAGACAGAGAGAACAAACAAGTTGCTTAGGGGACTCT	671
b	606		CTGCCATCAATATTACAAAGTATCCAGACAGAGAGAACAAACAAGTTGCTTAGGGGACTCT	665
y	672		GTCTTGATGCTCTCAATTGAACCTGCTGATCAAAAATGCTGATGGAAACTCAGGTTCCAAG	731
b	666		GTCTTGATGCTCTCAATTGAACCTGCTGATCAAAAATGCTGATGGAAACTCAGGTTCCAAG	725
y	732		AGTTTCTCAAGTGGCTCAACCCATCTTTCAACCTCTGAGAGAGAGTGGCCCTGGAGG	791
b	726		AGTTTCTCAAGTGGCTCAACCCATCTTTCAACCTCTGAGAGAGAGTGGCCCTGGAGG	785
y	792		ATGAAAACGATATGCAGATGGAGCTGAGACCGAGGTGGAATCTAAACCGCTGTCTGTGCCT	851
b	786		ATGAAAACGATATGCAGATGGAGCTGAGACCGAGGTGGAATCTAAACCGCTGTCTGTGCCT	845
y	852		GTGGAATTTGGGTCTGTACAGCCATGACCTGTGACCGAAGAGATCAAGAGGGGGCCAGA	911
b	846		GTGGAATTTGGGTCTGTACAGCCATGACCTGTGACCGAAGAGATCAAGAGGGGGCCAGA	905
y	912		CCCAGACAGAGGAGGAGATGACCAAGATATGTCAGGAGCTCCAAAAGCATCAGGAAACAG	971
b	906		CCCAGACAGAGGAGGAGATGACCAAGATATGTCAGGAGCTCCAAAAGCATCAGGAAACAG	965
y	972		CTGAAAAGACCAAGAGTGAACCAAGAGATCTAAATGAGGAGGCACAGACGATGC	1031
b	966		CTGAAAAGACCAAGAGTGAACCAAGAGATCTAAATGAGGAGGCACAGACGATGC	1025
y	1032		TGGATCCAGCATCTTCTCCACTTCAGCGCTGAGTTCAGTATACAAAGTGTCTGTCTACA	1091
b	1026		TGGATCCAGCATCTTCTCCACTTCAGCGCTGAGTTCAGTATACAAAGTGTCTGTCTACA	1085
y	1092		GTGCCAAATCACCAAGTATTGCTTATACCAATGAGTTTATTTTGTATTATTGTTTTT	1151
b	1086		GTGCCAAATCACCAAGTATTGCTTATACCAATGAGTTTATTTTGTATTATTGTTTTT	1145
y	1152		GCAATAAGGATATGAAGTGGCTGGCTAGGAAGGAGGCCACAGCTTCATTTCTAG	1211
b	1146		GCAATAAGGATATGAAGTGGCTGGCTAGGAAGGAGGCCACAGCTTCATTTCTAG	1205
y	1212		GAGTGCTTTAAGAGAACTGTAAATGTGCTCTGGGCTGGAGCTAGTAAGGAAACTGC	1271
b	1206		GAGTGCTTTAAGAGAACTGTAAATGTGCTCTGGGCTGGAGCTAGTAAGGAAACTGC	1265
y	1272		ATCACGATTGAAGAGGAACAGACCCAAATCTGAACCTCTTTTTCAGTTTACTGCACTGT	1331
b	1266		ATCACGATTGAAGAGGAACAGACCCAAATCTGAACCTCTTTTTCAGTTTACTGCACTGT	1325
y	1332		CACAGGCTCGAGGAGTGACACGATGCCAGAGAGAACTAGCAGAGGTGTCCCGGAGG	1391

Db	1326	CAGCAGGCTGCAGGAGTGCACACAGATGCAGAGAGAACTTAGCAGGGTGTCCCGGAGG	1386
Qy	1392	AGAGGTTTGGGAAAGCTCCACGGAGAGAAAGCTCTCTGTCTCCAGCCTCTTTTCCATTGCC	1451
Db	1386	AGAGGTTTGGGAAAGCTCCACGGAGAGAAAGCTCTCTGTCTCCAGCCTCTTTTCCATTGCC	1445
Qy	1452	GTACAGATGACAGACCTCCAGAGCATCCACGAGCATCTCTTTGGTCCCAATAACCTGCCTCTAGAT	1511
Db	1446	GTACAGATGACAGACCTCCAGAGCATCCACGAGCATCTCTTTGGTCCCAATAACCTGCCTCTAGAT	1505
Qy	1512	ACATAGCCATPACTGTCTAGTTTAAACCCAGTGTCCCTCAGACCTTGAGATGGAGTTTCTGGGAGG	1571
Db	1506	ACATAGCCATPACTGTCTAGTTTAAACCCAGTGTCCCTCAGACCTTGAGATGGAGTTTCTGGGAGG	1565
Qy	1572	GTACACCCAAATGATGCAGATCTTGTATCTTTGAGCCCTCTAGCGACCTAACACCAAAAT	1631
Db	1566	GTACACCCAAATGATGCAGATCTTGTATCTTTGAGCCCTCTAGCGACCTAACACCAAAAT	1625
Qy	1632	TTAAANAATCTTTTTTACCAGGTGCTATTCTCTGTAAACAC- TTTTCTTTGGCNAAGT	1690
Db	1626	TTAAANAATCTTTTTTACCAGGTGCTATTCTCTGTAAACACACTTTTTTTTGGCAAGT	1750
Qy	1691	TGACTTTATTTCTTCAATATTATCATATATATTATTTTAAATATTTTATTTTCTTGA	1745
Db	1686	TGACTTTATTTCTTCAATATTATCATATATATTATTTTAAATATTTTATTTTCTTGA	1745
Qy	1751	CTAGG 1755	
Db	1746	CTAGG 1750	
RESULT 11			
AK025860		1694 bp mRNA linear PRI 13-SEP-2003	
LOCUS		Homo sapiens cDNA: FLJ22207 fis, clone HRC01454.	
DEFINITION		AK025860	
ACCESSION		AK025860	
VERSION		AK025860.1 GI:10438502	
KEYWORDS		oligo capping; fis (full insert sequence).	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		1 Watanabe,X., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obyashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.	
TITLE		NEDO human cDNA sequencing project	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 1694)	
AUTHORS		Sugano,S., Suzuki,Y., Ota,T., Obyashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.	
TITLE		Direct Submission	
JOURNAL		Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:flcldna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)	
REFERENCE		NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).	
FEATURES		Location/Qualifiers	
source		1..1694	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="HRC01454"	
		/cell_type="primary human renal epithelial cells"	
		/clone_lib="HRC"	
		/note="cloning vector pME18SP13"	



[illegible]

841	AAAGGTTTTCACAGCCCTTGGACATCTGGGAGGAGGGCGAAGGTACACCAATTGTTAA	900
Db		
QY		
3024	AAGCAAGAAACACAGTGTCTCTTCCACTAGTCATTTAGAACATGGTTATCATCCAAGACT	3083
Db		
901	AAGCAAGAAACACAGTGTCTCTTCCACTAGTCATTTAGAACATGGTTATCATCCAAGACT	960
3084	ACTCTACCTGCAACATTGAACTCCAGACGAAATCCACATTCCTCTTGAGTTCTGCAG	3143
QY		
961	ACTCTACCTGCAACATTGAACTCCAGACGAAATCCACATTCCTCTTGAGTTCTGCAG	1020
Db		

3144	QY	CTTCGTGTAAATAGGGCAGCGTCTGCTTATATGCGCTAGAAATCACATGATCATCTGAGGACCAT	3203
1021	Db		
3204	QY	TCATGGAAGCTGCTAAATAGCCTAGCTCTGGGGAGTCTTCATATAAAGTTTTCATGAGCA	3263
1081	Db		
3264	QY	TATGGAAGCTGCTAAATAGCCTAGCTCTGGGGAGTCTTCATATAAAGTTTTCATGAGCA	3323
1141	Db		
3324	QY	TGACGGCTTCCTTGGGCTTTCTCTGTGTGTGTAGTTTGTATAAACAATATAGCATCTGTGA	3383
1201	Db		
3384	QY	AGATCCAGTGTCCATAGGAAACCTTCCCAATCCGCTGACTCTGGACTATATCAGTCTTTTG	3443
1261	Db		
3444	QY	GAAGCAGGGTTCTCTGCTCTCTAAAGCCACGTCGACAGTCTGAATGTCTTTCCT	3503
1321	Db		
3504	QY	TTACACCTATGTTTTTAAGTAGTCAAACTTCAAGAAACAATCTAAACAAGTTTCTCTGTGC	3563
1381	Db		
3564	QY	ATAATGTTTGTGAAGTCTGTATTTGTATTTAGTAGGCTCTATATGCAATTAACATTTGTT	3623
1441	Db		
3624	QY	TTTGTAACTCCGTGATCTTCTCTTTTCGGATATCTATTGTATGAATAAAGAAATTAAGTGAA	3683
1501	Db		
3684	QY	AAAAAATAA 3693	
1561	Db	AAAAAATAA 1570	
RESULT 13			
BC028921			
LOCUS	2851 bp mRNA linear ROD 07-OCT-2003		
DEFINITION	Mus musculus follistatin-like, mRNA (cdna clone MGC:25371		
ACCESSION	IMAGE:2647002), complete cds.		
VERSION	BC028921		
KEYWORDS	BC028921.1 GI:20810032		
SOURCE	MGC.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	1 (bases 1 to 2851)		
	Strausberg,R.L., Feingold,S., Grouse,L.H., Derge,J.G.,		
	Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,		

RESULT 13	
BC028921	
LOCUS	
DEFINITION	BC028921 2851 bp mRNA linear ROD 07-OCT-2003 Mus musculus follistatin-like, mRNA (cdna clone MGC:25371 IMAGE:3647002), complete cds.
ACCESSION	BC028921
VERSION	BC028921.1 GI:20810032
KEYWORDS	MGC.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2851)
REFERENCE	Klausner,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Straussberg,R.D., Collins,F.S., Wagner,L., Shennen,C.N., Schuler,G.D.,
AUTHORS	

Altechul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Uadin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Lequellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Farray,J., Helton,E., Kettenan,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smaluk,D.E., Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 2851)

Strausberg,R.

Direct Submission

Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)

Guarantee, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAP Plate: 30 Row: a Column: 7.

Location/Qualifiers

1. .2851

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="MGC:25371 IMAGE:2647002"

/tissue\_type="mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."

/clone\_lib="NCI CGAP Mam1"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. .2851

/gene="Pst1"

/note="synonym: TSC-36"

/db\_xref="LocusID:14314"

/db\_xref="GI:20810033"

/db\_xref="LocusID:14314"

/translation="MKRWLALSLVTIALVHGEERPSKSKI CANVFCAGRECAVTE KGEPTCLCIEQKPKRVCVSGNGKTYLNHICELHEDACLTGSKIQVDYDGHCKEKSA SPASPVVYQANDELRIQLEAEIIPDGFWSKSNYSLEILDYFKSPFNGKSH LDSSEFLKVEQNEQTAINITTYAQENNKILSLICVDALLIELSDENADWLKLSQEFLEK

CLNPSNPPEKKCALEDETYADGAETEVDNCRVCSCGHVYCTAMTCDDKQKGVQTH  
TEEEKTVGVQELQKHQGTAEKTKVNTKEI"

236. .370

/note="KAZAL; Region: Kazal type serine protease inhibitors"

/db\_xref="CDD:smart00280"

ORIGIN

Query Match 28.4%; Score 1049; DB 10; Length 2851;  
Best Local Similarity 76.9%; Pred No. 2,3e-247;  
Matches 1591; Conservative 0; Mismatches 390; Indels 88; Gaps 22;

QY 12 CTCCACCTTCGCTTACAGCTCGTCGCGCGCTCTCTGCCCGCGCCGCCAGAGACCTGG 71  
Db 12 CTCCACCTTCGCTTAACTCGCTGCGCGCACCTTGCCTCCGAGTGTCTCCGAGTCCCG 71

QY 72 ACCGACACCATGTGGAAACGCTCGCTCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCT 131  
Db 72 ACCGACACCATGTGGAAACGCTG-----CTGCGCTCTCGCTGCTGCTGCTGCTGCT 125

QY 132 GGGTCGCGCGCGAGGAGCTAAGGACAAATCCAAATCTGTGCCAAATGTGTTTGTG 191  
Db 126 TGGTCCACGCGCGAGGAGAACTTAGAAGCAATCCAAATCTGCGCCAAATGTGTTTGTG 185

QY 192 GAGCGCGCGGGAATGTGCAGTCACAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 251  
Db 186 GAGCTGGCAGGGAATGTGCCGTGTCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 245

QY 252 AATGCAAACTCACAAGAGGCGCTGTGTGGCAGTAATGGCAAGACCTACCTCAACCACT 311  
Db 246 AATGCAAACTCACAAGAGGCGCTGTGTGGCAGTAATGGCAAGACCTACCTCAACCACT 305

QY 312 GTGAATGATCGAGTGCCTGCTCAGTCCAGTCCAAATCCAAATCCAAATCCAAATCCAAAT 371  
Db 306 GTGAATTCATAGATGCTGCTGCTCAGTCCAGTCCAAATCCAAATCCAAATCCAAATCC 365

QY 372 ACTGCAAAAGAGAGAAATCCGTAAGTCCATCTGCCAGCGCCAGTCTGTTGCTATCAGTCCA 431  
Db 366 ACTGCAAAAGAGAGAGTCTGCGAGTCCATCTGCCAGCGCCAGTCTGTTGCTATCAAGCTA 425

QY 432 ACGGTGATGAGTCCGAGTGCATCTCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 491  
Db 426 ACGCGATGAGTCCGAGGCGGCGCTCATCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 485

QY 492 GCTGGTCTCTAAAGGAGCACTACAGTGAATCCTAGACAAATCTTAAAGAACTTTG 551  
Db 486 GCTGGTCTCTAAAGGAGCACTACAGTGAATCCTAGACAAATCTTAAAGAACTTTG 545

QY 552 ATATGCTGATCTCGCTGCGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 611  
Db 546 ATATGCGGCACTCTCACCTGGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 605

QY 612 CTGCCATCAATATTACAGTATCCAGACCCAGGAGAAACAAAGTTCGTAGGGGACTCT 671  
Db 606 CAGCCATCAACATCACCATTATGAGATCAGGAGAAACAAAGTTCGTAGGGGACTCT 665

QY 672 GTGTTGATGCTCTCATTTGAATGCTGTGTGATAAATGCTGATTTGGAATCTCAGCTTCCAAG 731  
Db 666 GTGTTGACGCGCTCATTTGAATGCTGTGTGATAAATGCTGATTTGGAATCTCAGCTTCCAAG 725

QY 732 AGTTTCTCAAGTCCCTCAACCCATCTTTCAACCTCTCAGGAGAAAGTGTGCTCTGGAGG 791  
Db 726 AGTTTCTCAAGTCCCTCAACCCATCTTTCAACCTCTCAGGAGAAAGTGTGCTCTGGAGG 785

QY 792 ATGAAACGTATCAGATGGAGTGGAGACCGAGGTGAGATGTAAACCGCTGTGTGTGCTGCT 851  
Db 786 AGCAAACTTATCAGATGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 845

QY 852 GTGGAATTTGGTCTGTACAGCCATCAGCTGTGAGGAGAAATCAGAGGGGGGGCCAG 911  
Db 846 GTGGCCACTTGGTCTGTACAGCAATGACCTGTGAGGAGAAATCAGAGGGGGGGCCAG 905

QY 912 CCCAGACAGAGGAGGAGATGACCATATGTCCAGAGCTCCAAAGATCATGAGAAACAG 971

906 CCCACAGAGGAGGAGAGACAGGATATGCTCAGGAATCTCCAGAGACACAGGACAG 965  
972 CTGGAAGACCAAGAGAGTGAGCACCAGAGATCTAATGAG-GAGGCAACAGACAGTCT 1030  
966 CAGAAAGACCAAGAGAGTGAAACCAACAGAGATCTAAGAGAGGACAGACACCGTGT 1025  
1031 CTGATCCAGCATCTCTCCATCTTCACATTCAGCGCTGAGTTTCAGTATACACAAAGTGTCTGTAC 1090  
1026 CCGGAGCCAGCGCT--CCTCTTCAGCGCTGAGCCAGTACACAC-AGAGTCTGCAG 1081  
1091 AGTCGCCAATCACCAGTATTTGCTTATATAGCAATGAGTTTATTTTATTTTATTTT 1150  
1082 AATCACAAATCTACTAGTATTTGCTGTATGCGAGCAATCTTATTG---TTTGT 1137  
1151 TGCATTAAGGATATGAAGTGGCTGGCTAGGAAGGAGGACAGCGCTTCATTCTTA 1210  
1138 TGCAATAAGGAATAGAGGTGGCCAGCTAGCGAGGAGGACACAACTTCACCTGTA 1197  
1211 GGAAGTCTTTAAGGAACCTGTAAATGCTGCTGGGCTGGAGGCTAGTAAGGAACCTG 1270  
1198 GGAATGCTTTAAGGAAC--TAAAGACACCTTGGGACGAGAGGCACTAAGGAACAG 1255  
1271 CATCACCATTAAGAGGACAGACACCAATCTGAAC---CTCTTTGAGTTTACTGCA 1326  
1256 CATC-GGTTGGCAGAGAGAGAGGAGGAGCTTTGAATGAAGCCTTCTGGGTCACAGCA 1314  
1327 TCTGTAGAGGCTGCGAGGAGTGCAACAGATGCCAGAGAACTTAGAGGGTGTCCCG 1386  
1315 GCTCGAGGAGATACAGGAAGAGCAT-----AGAGAAACATTGAACCTAGCGCTG 1364  
1387 GGAGGAGAGTTTGGGAGCTCCAGGAGAGGAGCGCTCTCTGCTCCAGCGCTCTTCCA 1446  
1365 CTGAGAGAGTGGGGAGCTTTGTAGGAGAGAACCTGCTGCTT---GACCTTGTCA 1421  
1447 TTGCCGTGAGATGACAGACTCCAGCATCCAGCATCTCTGGTCCCAATAC-TGCC 1505  
1422 CCACTGTGAGATGACAGACTGCGAGCAAGTCTCTCTCTCTCTGTCCTCCAAACATCAC 1481  
1506 CTAGATACATAGCCATCTGCTAGTTAAACCGAGTGCTCCCTCAGACTTGGAGTCTTCT 1565  
1482 CTGAAACACAGCGCCCACTAGTTA--CTGTGTCCTCAGCCTTGCATGAGTCTCC 1539  
1566 GGGAGGTACACCCAAATGATGCAATGATCTGTATATCTTTGAGCCCTTAGCGACCTAAC 1625  
1540 TGGAGGAGGTGTTTAAATGATGCAACATCTATGATCTCAAGCGCATGAGACCTAAC 1599  
1626 CAATTTTAAATACCTTTT-----ACCAAGGTGCTATTTC 1663  
1600 CAATTTTAAATACATTTTCTTTTTTTTTTTTGTAAACCAAGGTGCTATTTC 1659  
1664 TCTGTAACACATTTTTTTTGGCAAGTTGACTTTATCTTCAATTTATCATTAATTA 1723  
1660 TCTGTAACAGACTTTTT---CCAAGCTGACTTCAATCTCAGTTATTAACGTTATTA 1716  
1724 TTGT--TTTTTAATTTATTTTCTGACTAGTATTAAGCTTTGTAATTA-TTTTTT 1780  
1717 TTGTTGTTTTTAATTTTCAATTTTGTGACTAGATATTAAGCTTTTGTAAATTTTTC 1776  
1781 AGTAGTCCCACTTTCATAGTGGG-----AGGAGTTTGGGGTTCTTCTCTGTGCGAG 1833  
1777 ATTAGTCTACTATTTTCAAGAGTGAAGGTGAAGGGGTTTGGCAATTTTTCAGGTACAG 1836  
1834 GGGCTGAATTAACCCAGA-TGCTCCCACTTCCAGATACATAGTA-----TGAGCCCAT 1887  
1837 GGAATCTGTAAACCAACAGCCCATACCTCTGTACATATTAGACCGGTTGCGAGTTCGA 1896  
1888 GTTGGCC-----CCCTTAGCTTCCAGAGTCCA---CTATCTCCAGAGGAGCAAGGGTG 1939  
1897 GCTGCAACCCCAACCCAGAGCTTCTAGAAATCAGTCTCAATGCCAGAGGACAGAGGCC 1956  
1940 CCTTAGACCTTAAGCCAGGGAAGAGCATCTTCAATAAAACCTTCAAGATCCAAACA-T 1998

Db 1957 CCTCAGACAGAACCCACAGGACAAAGCATCTTCTATAGACAGCTCTTGAGATCCAAACAGT 2016  
QY 1999 TAATTTGTTTTTATTTATTTCTGAGAAGTT 2027  
Db 2017 TAATTTGTTTTTCTTCTGTAAGAGTT 2045  
MUSTSC36A 2823 bp mRNA linear ROD 27-OCT-1993  
DEFINITION Mus musculus TGF-beta-inducible protein (TSC-36) mRNA, complete cds.  
ACCESSION M91380  
VERSION M91380.1 GI:349005  
KEYWORDS TGF-beta-inducible protein.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 2823)  
AUTHORS Shibanuma,M., Mashimo,J., Mita,A., Kuroki,T. and Nose,K.  
TITLE Cloning from a mouse osteoblastic cell line of a set of transforming-growth-factor-beta 1-regulated genes, one of which seems to encode a follistatin-related polypeptide  
JOURNAL Eur. J. Biochem. 217 (1), 13-19 (1993)  
MEDLINE 94039028  
PUBMED 7901004  
REFERENCE 2 (bases 1 to 2823)  
AUTHORS Nose,K.  
TITLE Direct Submission  
JOURNAL Submission received (Apr 2 1992) via NEC Floppy by Genbank (1992)  
COMMENT Original source text: Mus musculus cDNA to mRNA.  
FEATURES  
Location/Qualifiers  
1..2823  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/cell\_lines="MC3T3-E1"  
/cell\_type="osteoblast"  
1..2823  
/gene="TSC-36"  
80..1000  
/gene="TSC-36"  
/codon\_start=1  
/product="TGF-beta-inducible protein"  
/protein\_id="AAC37633.1"  
/db\_xref="GI:349005"  
/translation="MWKRLALSLVTIALVHGEEPRSKIKANVFCGAGRECAVTE KGEPTCLISQCKPKKPVCGSNKTYLNHCELHRACTLGSKTQVDYDGHCKEKKA SPSASPVVQANDELRRLLIOWLEAIIIPDGFSGSNYSLELDKXFKSPDNGDSH LDSSEFLKVFQNETAINITTYADENNKLLRSLCDALLIELSDENADWKLSPFELK CLNPSFNPEKKCALEVETVADGAETEDCNRCVSCGHWVCTAMTCDGNKQGVQTH TEEBKTYVBLQKHQGTAEKTKVNTKEI"  
ORIGIN  
Query Match 27.8%; Score 1025.4; DB 10; Length 2823;  
Best Local Similarity 75.9%; Pred. No. 1.5e-241; Indels 93; Gaps 21;  
Matches 1572; Conservative 0; Mismatches 406;  
QY 12 CTCACACCTCCGCTTACAGCTCGCTGCGCGCTCTGCTCCCGCCCTGCTCCCGAGGACCTGG 71  
Db 9 CTCACACCTTCGCTCTAACTCGCTGCGCCACCTGCTCCGAGTCTCTCCGAGTCCCG 68  
QY 72 ACCAGACACCATGTGGAAGCTGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 131  
Db 69 ACCGAGACCATGTGGAACGATGG-----CTGGCGCTCTCGCTGGTACCATCGGCC 122  
QY 132 GGGTCCGCGCGAGGAAGAGTAAGAGGACAAATCAAGATCTGTGCCAATGTGTTTGTG 191  
Db 123 TGTCCACGCGAGGAGGAACTAGAGCAATCAAGATCTGCGCAATGTGTTTGTG 182  
QY 192 GAGCGCGCGGGAATGTGCACTCAGAGAAAGGGAAACCCACTGTCTCTGCAATTGAGC 251

Db 183 GAGCTGGCAGGGAATGTGCGGTACAGAGAGGGGAGCCCAAGTGCCTCTGCAATTGAGC 242  
2y 252 AATGCAAACTCACAAGAGCGCTGTGTGTGGAGATATGCAAGACATCCTCAACCACT 311  
Db 243 AATGCAAACTCACAAGAGCGCTGTGTGTGGAGATATGCAAGACATCCTCAACCACT 302  
2y 312 GTGAATGCAATCGAGATGCTGCTGCTCACTGATCCAAATCCAGAGTTGATTAGATGGAC 371  
Db 303 GTGAATTCATAGAGATGCTGCTGCTCACTGATCCAAATCCAGAGTTGATTAGATGGAC 362  
2y 372 ACTGCAAGAGAGAAATCCGTAACTAGTCCATCTGCGAGCCCGAGTTGTTGATATCAATGCA 431  
Db 363 ACTGCAAGAGAGAAATCCGTAACTAGTCCATCTGCGAGCCCGAGTTGTTGATATCAATGCA 422  
2y 432 ACCGTGATGAGCTCCGAGCTGCGATCATCAGTGGCTGGAAGCTGAGATCATCTCAGATG 491  
Db 423 ACCGATGAGCTCCGAGCTGCGATCATCAGTGGCTGGAAGCTGAGATCATCTCAGATG 482  
2y 492 GCTGGTCTCTAAAGGAGCAATACAGTGAATCTTAGCAAGATATTTAAGAACTTTG 551  
Db 483 GCTGGTCTCTAAAGGAGCAATACAGTGAATCTTAGCAAGATATTTAAGAACTTTG 542  
2y 552 ATATGATGATCTCGCTGAGCTCCAGTGAATCTTGAAGTTTGTGGAACAGAAATGAA 611  
Db 543 ATATGAGGACTCTCAGCTGATCTCCAGTGAATCTTGAAGTTTGTGGAACAGAAATGAA 602  
2y 612 CTGCCATCAATATPACAACTGATCCAGACAGAGAAACAAAGTTGCTTAGGGGACTCT 671  
Db 603 CAGCCATCAATCAACCACTATGAGATCAGAGAGAAACAAAGTTGCTTAGGGGACTCT 662  
2y 672 GTGTTGATGCTCTATGAACTGCTGATGAATGCTGATGGAACACTCAGCTTCCAG 731  
Db 663 GTGTTGAGCCCTCATTTGAACTGCTGATGAGAACTGCTGGAACACTCAGCTTCCAG 722  
2y 732 AGTTTCTCAAGTGCCTCAACCCATCTTTCAACCTCTGAGAGAAAGTGTGCTCGAGG 791  
Db 723 AGTTTCTCAAGTGCCTCAACCCATCTTTCAACCTCTGAGAGAAAGTGTGCTCGAGG 782  
2y 792 ATGAAAGTATGAGATGAGCTGAGACCGAGTGGACTGATACCCCTGTGTGTGCT 851  
Db 783 TCGAAAGCTATGAGATGAGCTGAGACTGAGGTGGACTGCAATGCTGTGTGTGCT 842  
2y 852 GTGAAATTTGGTCTGTACAGCATGACCTGTGACGGAAGAAATCAGAGGGGGCCAG 911  
Db 843 GTGSCCACTGGTCTGACAGCAATGACCTGTGATGGAAGAAATCAGAGGGGGTCCAG 902  
2y 912 CCCAGACAGAGGAGATGACAGATATGTCCAGAGCTCCAAAGCATCAGGAACAG 971  
Db 903 CCCACAGAGAGGAGAGAGATATGTCCAGAGAACTCCAGAGCAACAGGGCCAG 962  
2y 972 CTGAAAGACCAAGAGATGAGCACCACCAAGATCTAATGAG-GAGGCAAGACCAAGTGT 1030  
Db 963 CAGAAAGACCAAGAGATGAGCACCACCAAGATCTAAGAGAGGACAGACCGTGT 1022  
2y 1031 CTGATATCCAGCATCTTCTCCATTCAGCGCTGAGTTTCAATATACAGATGCTGCTAC 1090  
Db 1023 CCGAGAGCCAGCGCT---CCTCTTACGCGCTGAGCCGAGTACACAC-AGAGTCTGAGC 1078  
2y 1091 AGTCGCCAAATCACAGTATTTGCTTATATAGCAATGAGTTTATTTTATTTTGT 1150  
Db 1079 AATCACCAATCATAGTATTTGCTTATATGCGAGCAATCTTATTTG---TTTGT 1134  
2y 1151 TGCAATAAGATATGAGTGTGCTGCTAGGAAGGAGGCGCACAGCTTCAATTTCTA 1210  
Db 1135 TGCAATAAGGAAATGAGGTGGCGAGCTAGCGAGGGAAGGCGCACAACTTCCCTGT 1194  
2y 1211 GGAGTCTTTAAGAGAACTGTAATGTGCTGCTGGGCTGGAGGCTAGTAAGAAACTG 1270  
Db 1195 GGAATGCTTTAAGAGAACT---TAAAGACACTTGGACGAGAGGCACTAAGAAACAG 1252  
2y 1271 CATCAGATTGAAGAGAGAACAGACCCAAATCTGAAC---CTCTTTTGTAGTTTACTGCA 1326  
Db 1253 CATC-GGTTGGCAGAGAGAGAGGAGGAGTGAATGAAGCCCTTCTGGGTCACAGCA 1311

Qy 1327 TCTGTACAGCGCTGCAGGAGTGCACACAGTCCAGAGAGAACTTAGCAGGTGTGCCC 1386  
Db 1312 GCTCGAGGAGAAATACAGAAAAGCAT-----AGAGAAACATTTAGACTAGCCCTG 1361  
Qy 1387 GGAGAGAGAGTTTGGGAGCTCCACGGAGAGAAAGCTCTCTGCTTCCAGCCTCTTTCCA 1446  
Db 1362 CTGAGGAGAGTGGGGGAGCTTTGTAGGAGGAACCTGCTGCTTT---GACCCCTGTCA 1418  
Qy 1447 TTGCGCTCAGATGACAGACCTCCAGCATCCAGCATCTCTTGTGTCCTCAATAC-TGCT 1505  
Db 1419 CCATGCTCAGATGACAGACCTCCAGCAAGTCTCTCTCTCTTTTGTGTCCTCAATAC 1478  
Qy 1506 CTAGATACATAGCCATCTGCTAGTAAACCCAGTGTCCCTCAGACTTGGATGGAGTTCT 1565  
Db 1479 CTGAACACAGCGCCCAACTAGTACCTGTGT---CCTCAGCCTTGCATGGAGTTCC 1535  
Qy 1566 GGGAGGTACACCCAAATGATGAGATCTTGTATATCTTTGAGCCCTTAGCAGCTTAAC 1625  
Db 1536 TGGAGGAGGTGTTTAAATGATGACAGACTTATGACTTCAAGCGCAT--GGAGACTAAC 1593  
Qy 1626 CAAATTTTAAATACTTTTT-----ACCAAAGGTGCTATTTT 1663  
Db 1594 CAAATTTTAAATACTTTTTTTTTTTTTTTTTTTTGTAAACCAAGGTGCTATTT 1653  
Qy 1664 TCTGTAACACTTTTTTTTGGGAGTGTGCTTATTTCTTCAATATATCAATATTA 1723  
Db 1654 TCTGTAACAGCTTTTTT---CAAGCTGACTTCAATCTCAGTATTTACCGTTATATTA 1710  
Qy 1724 TGT--TTTTTAATATTTTATTTTCTGACTAGTATTAAGCTTTTGTATATTTTCA 1781  
Db 1711 TGTGTTTTTTTAAATATTTTCAATTTTGTAGTATATTAAGCTTTTGTATATTTTCA 1770  
Qy 1782 GTAGTCCACCACTTC-----ATAGTGAAGAGTGTGGGTCTCTCTGCTGCA-- 1832  
Db 1771 TTAGTCTTACTATTTTCGAGAAAGTGAAGTGAAGGGGTGGGCAATTTTCCAGGGTACA 1830  
Qy 1833 -GGGCTCAAAATACCCAGATGCTCCACCTGCGCACATACATAG---TGCAGCCCAT 1886  
Db 1831 GGGAACTCTGTAACAAAGAGCCCATACCTCTGACATATAGACGGTGTGAGTTCGG 1890  
Qy 1887 AGTTGGCCCCCTAGCTTCCAGAGTCCATATCTGC-----CAGAGGAGCAAGG 1937  
Db 1891 AGCATGACCCCAACCCAGAGCTTCTAGAAAATCAGCTCCATGCGCAGGAAGCAAGAG 1950  
Qy 1938 TGCCTTAGACTAGCCAGGGGAGAGAGCATCTTCAATAAAACATTTCAAGATCCAAACA 1997  
Db 1951 GCCCTCAGCAGAGGCCACAGCAAGAGCATCTCTATAGACGCTGTGAGATCCAAACA 2010  
Qy 1998 -TTAATTTTGTTTTATTTTATTTCTGAGAAATT 2027  
Db 2011 GTTAATTTGCTTTTGTCTTCTGTAAGAAGTT 2041

RESULT 15  
HSM800525  
LOCUS  
DEFINITION Homo sapiens mRNA; cdna DXFp5860223 (from clone DKFp5860223).  
ACCESSION AL050228  
VERSION AL050228.1 GI:4884472  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1010)  
AUTHORS Anseorge, W., Winkler, U., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
TITLE Direct Submission  
JOURNAL Submitted (15-MAY-1999) MIPS, Am Klopferspitz 18a, D-82152  
Martinsried, GERMANY  
COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by EMBL (European Molecular Biology Laboratories,



Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.  
This clone (DKFZp5860223) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

# FEATURES

source 1..1010  
    /organism="Homo sapiens"  
    /mol\_type="mRNA"  
    /db\_xref="RZPD:DKFZp5860223"  
    /db\_xref="taxon:9606"  
    /clone="DKFZp5860223"  
    /tissue\_type="uterus"  
    /clone\_lib="586 (synonym: hutel). Vector pSport1; host DH10B; sites NorI + SalI/Wlul"  
    /dev\_stage="adult"  
polyA\_signal 991..996  
polyA\_site 999  
RIGIN

Query Match 27.3%; Score 1008.4; DB 9; Length 1010;  
Best Local Similarity 99.9%; Pred. No. 1.9e-237;  
Matches 1009; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2683	TTGACATAAGATCTCATTTGACAGAAAGCAGATTAAAGACCATCAGAGAAATTTATTTAG	2742
1	TTGACATAAGATCTCATTTGACAGAAAGCAGATTAAAGACCATCAGAGAAATTTATTTAG	60
2743	GTGTGAATGACAGAGCACTGTGAGAACTGTGTGCGCAAAATAGAAATTCCTCTAGTT	2802
61	GTGTGAATGACAGAGCACTGTGAGAACTGTGTGCGCAAAATAGAAATTCCTCTAGTT	120
2803	TTTCTTGTCTCATTTGAAAGGAGAAATTCACCTTTGTTAGCATTTCAAGCTTTTATG	2862
121	TTTCTTGTCTCATTTGAAAGGAGAAATTCACCTTTGTTAGCATTTCAAGCTTTTATG	180
2863	TATCCATCCCATCTAAATCTTCAACTTCCACTTGTTCAGTCTGAAATGCAGTCCCT	2922
181	TATCCATCCCATCTAAATCTTCAACTTCCACTTGTTCAGTCTGAAATGCAGTCCCT	240
2923	GTCCAAAGTGCTTTGGAGAACTCAGCAGCAGCAGCCCTTAATCAAGGTTTTTACGCGCTT	2982
241	GTCCAAAGTGCTTTGGAGAACTCAGCAGCAGCAGCCCTTAATCAAGGTTTTTACGCGCTT	300
2983	GGACATATCGGAGGAGGCAAGAGTACACCAATTTGTTAAAGCAAGAACCAAGTGT	3042
301	GGACATATCGGAGGAGGCAAGAGTACACCAATTTGTTAAAGCAAGAACCAAGTGT	360
3043	CTCTTCACTAGTCAATTAGAACATGGTTATCATCCAGACTACTCTACCTGCAACATTG	3102
361	CTCTTCACTAGTCAATTAGAACATGGTTATCATCCAGACTACTCTACCTGCAACATTG	420
3103	AATCCCAAGAGCAATCCACATTCCTCTTGGATTTCTGAGCTTCTGTAAATAGGGCA	3162
421	AATCCCAAGAGCAATCCACATTCCTCTTGGATTTCTGAGCTTCTGTAAATAGGGCA	480
3163	GCTGTGCTCTATGCGGTAGAAATCACATGATCTCAGGACCAATTCATGAAAGCTGCTAAATA	3222
481	GCTGTGCTCTATGCGGTAGAAATCACATGATCTCAGGACCAATTCATGAAAGCTGCTAAATA	540
3223	GCCTAGTCTGGGAGTCTTCCATAAAAGTTTGGCATGGAGCAAAACAAACAGGATTAACCTA	3282
541	GCCTAGTCTGGGAGTCTTCCATAAAAGTTTGGCATGGAGCAAAACAAACAGGATTAACCTA	600
3283	GGTTTGGTCTCTCAGCCCTCTAAAGCATAGGCTTTAGCTCAGGCTTCTCTTGGGCTT	3342
601	GGTTTGGTCTCTCAGCCCTCTAAAGCATAGGCTTTAGCTCAGGCTTCTCTTGGGCTT	660
3343	TCTCTGTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAA	3402
661	TCTCTGTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAA	720

Search completed: March 11, 2004, 03:35:56  
Job time : 9420.52 secs

Qy	3403	ACCTTCCACATGCGTGACTCTGGACTATATCAGTTTGTGAAAGCAGGGTTCCTCTGC	3462
Db	721	ACATTCACATGCGTGACTCTGGACTATATCAGTTTGTGAAAGCAGGGTTCCTCTGC	780
Qy	3463	CTGTAAACAAGCCACACGTGACAGCTCTGAATGCTTTCCCTTTACACCTATGTTTTTAAG	3522
Db	781	CTGTAAACAAGCCACACGTGACAGCTCTGAATGCTTTCCCTTTACACCTATGTTTTTAAG	840
Qy	3523	TAGTCAAACTTCAAGAAACAATCTAAACAAGTTTCTGTTGCATATGTTTGTGAACCTG	3582
Db	841	TAGTCAAACTTCAAGAAACAATCTAAACAAGTTTCTGTTGCATATGTTTGTGAACCTG	900
Qy	3583	TATTTGTATTAGTAGGCTTCTATATTGCAATTTAACTTTGTTTGTAACTCCTGATTCCT	3642
Db	901	TATTTGTATTAGTAGGCTTCTATATTGCAATTTAACTTTGTTTGTAACTCCTGATTCCT	960
Qy	3643	CCTTTTCGGATCTATTGATGAATTAAGAAATTAAGTGAAGAAAAA	3692
Db	961	CCTTTTCGGATCTATTGATGAATTAAGAAATTAAGTGAAGAAAAA	1010



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

M nucleic - nucleic search, using sw model

un on: March 10, 2004, 08:01:57 ; Search time 898.804 Seconds  
(without alignments)  
17454.974 Million cell updates/sec

itle: US-10-084-817-20

effect score: 3693  
sequence: 1 Gcagcgaggactccacact.....ttaaagtgaataaaaaaaa 3693

coring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

earched: 3373863 seqs, 2124099041 residues

otal number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase : N\_Geneseq\_29Jan04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002s:\*
- 7: geneseqn2003as:\*
- 8: geneseqn2003bs:\*
- 9: geneseqn2003cs:\*
- 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

result No.	Score	Match	Length	DB	ID	Description
1	3693	100.0	3693	10	ADE77194	Ade77194 Human cDN
2	1940	52.5	1987	6	ABK64722	Abk64722 Human ben
3	1940	52.5	1987	6	ABN96790	Abn96790 Gene #328
4	1729.8	46.8	2538	7	AB234719	Ab234719 Coding se
5	1729.8	46.8	2538	7	AA157940	AA157940 Human pol
6	1678.4	45.4	1755	4	AA137538	AA137538 Human mus
7	1678.4	45.4	1755	4	ABX60526	ABX60526 CDNA enco
8	1008.4	27.3	1010	6	ABQ88210	ABQ88210 Human bst
9	924.2	25.0	1017	2	AA139840	AA139840 Gastric c
10	911.8	24.7	926	2	AA168831	AA168831 cDNA for
11	850.2	23.0	1370	9	ADBS8526	ADBS8526 Toxicity-
12	844	22.9	879	6	AB929899	AB929899 Human cDN
13	757.2	20.5	881	2	AA139842	AA139842 Gastric c
14	729.4	19.8	930	7	AB234644	AB234644 Coding se
15	729.2	19.7	810	2	AA139843	AA139843 Gastric c
16	674.8	18.3	808	2	AA139844	AA139844 Gastric c
17	658.4	17.8	879	2	AA139807	AA139807 Gastric c
18	657.2	17.8	705	6	AB929891	AB929891 Human cDN
19	657	17.8	719	2	AA139803	AA139803 Gastric c
20	630.4	17.1	634	4	AA123550	AA123550 Human lun
21	630.4	17.1	634	9	ADBS6848	ADBS6848 Human lun
22	630.4	17.1	634	9	ADBS68102	ADBS68102 Human lun
23	629.4	17.0	708	2	AA139805	AA139805 Gastric c

## ALIGNMENTS

### RESULT 1

ADe77194  
ID ADE77194 standard; cDNA; 3693 BP.

XX AC ADE77194;

XX DT 29-JAN-2004 (first entry)

XX DE Human cDNA differentially expressed in a liver disorder #271.

XX KW human; ss; gene; liver disorder; hyperlipidaemia; hypertension;

XX KW type II diabetes; tumour; liver; inflammatory disorder;

XX KW immune response disorder; high-throughput screening;

XX KW differential gene expression; gene therapy.

XX OS Homo sapiens.

XX FN US2003108871-A1.

XX PD 12-JUN-2003.

XX PF 30-JUL-2001; 2001US-00919039.

XX PR 28-JUL-2000; 2000US-0222113P.

XX PA (KASE/) KASER M R.

XX PI Kaser MR;

XX DR WPI; 2004-031227/03.

XX DR P-PSDB; ADE77195.

XX Composition comprising several cDNAs that are differentially expressed in treated human C3A liver cell cultures, useful for treating liver disorders.

XX PS Claim 1; SEQ ID NO 359; 41pp; English.

XX The invention relates to a composition comprising several cDNAs that are differentially expressed in a liver disorder. The composition is useful for treating liver disorder such as hyperlipidaemia, hypertension, type II diabetes, tumours of the liver and disorders of the inflammatory and immune response. The composition is useful for a high-throughput method of screening several molecules or compounds to identify a ligand which specifically binds a cDNA. A protein encoded by the cDNA is useful for a high-throughput method for using a protein to screen several molecules or



1921 TGCCAGAGGAGCAAGGGTGCCTTAGACCTAAGCCAGGGAAGAGCATCTTCATAAAAA 1980  
 1981 CTTTCAAGATCCAAACATTAATTTGTTTTATTTATTTCTGAGAGTTGAGGCAAAATCAGT 2040  
 1981 CTTTCAAGATCCAAACATTAATTTGTTTTATTTATTTCTGAGAGTTGAGGCAAAATCAGT 2040  
 2041 ATTCCCAAGGATGGCGACCAAGGCGACCAAGCAGGCTTAGGATATCCAGGCTTACCAAT 2100  
 2041 ATTCCCAAGGATGGCGACCAAGGCGACCAAGCAGGCTTAGGATATCCAGGCTTACCAAT 2100  
 2101 ATGCTCATTCGACTAACTAGAGGGGTGAGTTGGCCCTGCTCTCTCTCTCTCTCTCTCTCT 2160  
 2101 ATGCTCATTCGACTAACTAGAGGGGTGAGTTGGCCCTGCTCTCTCTCTCTCTCTCTCTCT 2160  
 2161 AGTTTCCTCAGTGGTGAAGATGCACTAATACCTTTTGAATTTGATTAAGTTATTAATTC 2220  
 2161 AGTTTCCTCAGTGGTGAAGATGCACTAATACCTTTTGAATTTGATTAAGTTATTAATTC 2220  
 2221 TGTGGTCTTGATCATTTGGTCCAGAGGGGAGATAGTTCCCTGATTTTTCCTTCTCTCTCT 2280  
 2221 TGTGGTCTTGATCATTTGGTCCAGAGGGGAGATAGTTCCCTGATTTTTCCTTCTCTCTCT 2280  
 2281 ATAGATAAATTAATTAATTTGTTACTAGAAACAAGAAATGTCAGATGGCCCAAAAAACAAGATG 2340  
 2281 ATAGATAAATTAATTTGTTACTAGAAACAAGAAATGTCAGATGGCCCAAAAAACAAGATG 2340  
 2341 ACCAGATTTGATCTCAGCTGATGACCTACAGGTGCTGATGATGAGTCCCTCATG 2400  
 2341 ACCAGATTTGATCTCAGCTGATGACCTACAGGTGCTGATGATGAGTCCCTCATG 2400  
 2401 GGTAAAGCAGGAGAGAGTGGGAAAGAGAACCCACCTCTCTCTCATTTTTCATTT 2460  
 2401 GGTAAAGCAGGAGAGAGTGGGAAAGAGAACCCACCTCTCTCTCATTTTTCATTT 2460  
 2461 CATGTTAACTTCGGCTGGAATAGAAAGCATTCCTTAGAGATGAGGATAAAGAAAG 2520  
 2461 CATGTTAACTTCGGCTGGAATAGAAAGCATTCCTTAGAGATGAGGATAAAGAAAG 2520  
 2521 TTTTCAGATTTCAACAGGGGGAGAAATGGAGATTTAATCCTAAATCTGACCTGGGGAG 2580  
 2521 TTTTCAGATTTCAACAGGGGGAGAAATGGAGATTTAATCCTAAATCTGACCTGGGGAG 2580  
 2581 GTACGTCATTTACAGTTAGTCTGCTGCTTTCGACTCTGCTGATTAATACCCCACTCAC 2640  
 2581 GTACGTCATTTACAGTTAGTCTGCTGCTTTCGACTCTGCTGATTAATACCCCACTCAC 2640  
 2641 TACCCTGTTTCAGATGATTTGGAAATACCAAGATTAATCCTTGACATTAAGATCTCAT 2700  
 2641 TACCCTGTTTCAGATGATTTGGAAATACCAAGATTAATCCTTGACATTAAGATCTCAT 2700  
 2701 TGCAGAAAGCAGATTAAGACCATCAGAGGAAATTTATTTAGTGTGTAATGCAAGGCAA 2760  
 2701 TGCAGAAAGCAGATTAAGACCATCAGAGGAAATTTATTTAGTGTGTAATGCAAGGCAA 2760  
 2761 CTGTGAGAACTGTGTGCCAAAATAGAAATCCTCTGATTTTCTGTTCTCATTTGA 2820  
 2761 CTGTGAGAACTGTGTGCCAAAATAGAAATCCTCTGATTTTCTGTTCTCATTTGA 2820  
 2821 AAGGAGAAATTCACCTTTGTTAGCAATTCAGACTTTATGATCCATCCCATCTAAAA 2880  
 2821 AAGGAGAAATTCACCTTTGTTAGCAATTCAGACTTTATGATCCATCCCATCTAAAA 2880  
 2881 ACTCTTCAAACTCCACTGTTGCTGAGATGAGATGAGTCCCTGTCAGGTCCTTGAGAG 2940  
 2881 ACTCTTCAAACTCCACTGTTGCTGAGATGAGATGAGTCCCTGTCAGGTCCTTGAGAG 2940  
 2941 ACTCACAGCAGCCTTAACTAAAGTTTTTACAGCCCTTGGACACTATGGAGGAGG 3000  
 2941 ACTCACAGCAGCCTTAACTAAAGTTTTTACAGCCCTTGGACACTATGGAGGAGG 3000  
 3001 GCAAGAGTACCAATTTGTTTAAAGCAAGAAACCAAGTGTCTCTCTCACTAGTCATTTA 3060

3001 GCAAGAGTACCAATTTGTTTAAAGCAAGAAACCAAGTGTCTCTCTCACTAGTCATTTA 3060  
 3061 GAAACATGTTTATCATCCAGAGTACTCTACCTGCAACATTTGAATCCCAAGAGCAAAATC 3120  
 3061 GAAACATGTTTATCATCCAGAGTACTCTACCTGCAACATTTGAATCCCAAGAGCAAAATC 3120  
 3121 CACATTTCTCTGAGTTCTGCAGCTTCTGTGTAATAAGGCGAGTGTCTCTATGCCGTA 3180  
 3121 CACATTTCTCTGAGTTCTGCAGCTTCTGTGTAATAAGGCGAGTGTCTCTATGCCGTA 3180  
 3181 GAATCACATGATCTGAGGACCATTTATGGAAGTGTCTGCTAGTCTGCGGAGTCT 3240  
 3181 GAATCACATGATCTGAGGACCATTTATGGAAGTGTCTGCTAGTCTGCGGAGTCT 3240  
 3241 TCCATAAAGTTTTGCATGGAGCAAAACAAGGATTAACCTAGGTTTGGTTCCCTTCAGCC 3300  
 3241 TCCATAAAGTTTTGCATGGAGCAAAACAAGGATTAACCTAGGTTTGGTTCCCTTCAGCC 3300  
 3301 CTCTAAAAGCATAGGGCTTAGCCTGCAAGGCTTCTTGGGCTTTCTCTGTGTGTAGTTT 3360  
 3301 CTCTAAAAGCATAGGGCTTAGCCTGCAAGGCTTCTTGGGCTTTCTCTGTGTGTAGTTT 3360  
 3361 TGTAAACACTATATGATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACATGCCGTG 3420  
 3361 TGTAAACACTATATGATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACATGCCGTG 3420  
 3421 ACTCTGGACTATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTGCTTAAAGAGAAA 3480  
 3421 ACTCTGGACTATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTGCTTAAAGAGAAA 3480  
 3481 GGACAGTCTGAATGCTTCTTCTTACACCTATGTTTAAAGTGTCAAACTTCAAGAAA 3540  
 3481 GGACAGTCTGAATGCTTCTTCTTACACCTATGTTTAAAGTGTCAAACTTCAAGAAA 3540  
 3541 CAATCTAAACAAAGTTCTGTTGCAATATGTTTCTGAACTGTATTTGATTTAGTAGGC 3600  
 3541 CAATCTAAACAAAGTTCTGTTGCAATATGTTTCTGAACTGTATTTGATTTAGTAGGC 3600  
 3601 TCTATATGCAATTAACCTGTTTGTAACTCTGATTTCTCTTTTCGGATACTATTG 3660  
 3601 TCTATATGCAATTAACCTGTTTGTAACTCTGATTTCTCTTTTCGGATACTATTG 3660  
 3661 ATGATAAAGAAATTAAGTGAAGAAAAA 3693  
 3661 ATGATAAAGAAATTAAGTGAAGAAAAA 3693

RESULT 2

ABK64722

ID ABK64722 standard; DNA; 1987 BP.

XX ABK64722;

XX AC ABK64722;

XX DT 18-JUN-2002 (first entry)

XX Human benign prostatic hyperplasia gene #617.

XX Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.

XX Homo sapiens.

XX WO200212440-A2.

XX 14-FEB-2002.

XX 07-AUG-2001; 2001WO-US024708.

XX 07-AUG-2000; 2000US-0223323P.

XX 05-JUN-2001; 2001US-00873319.

XX (GENE-) GENE LOGIC INC.

XX (NIBS) JAPAN TOBACCO INC.

PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;  
CX WFI; 2002-257476/30.

CX Identifying drugs for and diagnosing benign prostatic hyperplasia, by  
CX detecting expression levels of one or more genes in prostate cells from  
CX patient that are differentially regulated compared to normal prostate  
CX cells.

CX Disclosure; Page 328-329; 444pp; English.

CX The invention relates to a method of diagnosing (I) the onset or  
CX progression of benign prostatic hyperplasia (BPH), or screening (II) for  
CX or identifying an agent that modulates the onset or progression of BPH.  
CX The method is based on changes in gene expression in BPH tissue isolated  
CX from patients exhibiting different clinical states of prostate  
CX hyperplasia as compared to normal prostate tissue. (I) comprises  
CX detecting the expression levels of one or more genes in prostate cells  
CX from the subject that are differentially regulated compared to normal  
CX prostate cells. (II) comprises preparing a first gene expression profile  
CX of BPH cells or BPH-like cell population, exposing the cells to the  
CX agent, preparing a second gene expression profile of the agent exposed  
CX cells, and comparing the first and second gene expression profiles. (I)  
CX is useful for diagnosing the onset or progression of BPH. (II) is useful  
CX for identifying an agent that modulates the onset or progression of BPH.  
CX The methods are useful to present information identifying the expression  
CX level in a tissue or cells, by comparing the expression level of genes  
CX given in the specification in the tissue or cells to the level of  
CX expression of gene in the database, and displaying the expression levels  
CX of at least one gene in the tissue or cell sample compared to the  
CX expression level in BPH. Agents using (II) are useful for treating BPH or  
CX prostate cancer. ABK64106-ABK64960 represent human benign prostatic  
CX hyperplasia gene sequences of the invention

CX Sequence 1987 BP; 522 A; 482 C; 498 G; 485 T; 0 U; 0 Other;

CX Query Match 52.5%; Score 1940; DB 6; Length 1987;  
CX Best Local Similarity 99.7%; Pred. No. 0;  
CX Matches 1954; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

2Y 12 CTCACCTCCGCTTACAGTCGTCGCGCGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 71  
2b 21 CTCACCTCCGCTTACAGTCGTCGCGCGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 80  
2Y 72 ACCAGACCGAGTGGAAAGCTGGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 131  
2b 81 ACCAGACCGAGTGGAAAGCTGGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 140  
2Y 132 GGGTCCGCGCGGAGAGCTAAGGAGCAATCCAAGATCTGTGCCAATGTGTTTGTG 191  
2b 141 GGGTCCGCGCGGAGAGCTAAGGAGCAATCCAAGATCTGTGCCAATGTGTTTGTG 200  
2Y 192 GAGCGCGCGGAGTGCAGTCACAGAGAGGGAACCCACCTCTCTCTCATTTGAGC 251  
2b 201 GAGCGCGCGGAGTGCAGTCACAGAGAGGGAACCCACCTCTCTCTCATTTGAGC 260  
2Y 252 AATGCAAACTCACAAGAGCGCTGTGTGGCAGTAAATGCAAGACCTACCTCAACCACT 311  
2b 261 AATGCAAACTCACAAGAGCGCTGTGTGGCAGTAAATGCAAGACCTACCTCAACCACT 320  
2Y 312 GTGACCTGCAATCGAGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 371  
2b 321 GTGACCTGCAATCGAGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 380  
2Y 372 ACTGCAAGAGAGAAATCCGTAAGTCCATCTGCGCAGCCAGTTGTTGCTATCAGTCCA 431  
2b 381 ACTGCAAGAGAGAAATCCGTAAGTCCATCTGCGCAGCCAGTTGTTGCTATCAGTCCA 440  
2Y 432 ACCGTGATGAGCTCCGAGTCGATCATCCAGTGGCTGGAGCTGAGATCATCTCAGATG 491  
2b 441 ACCGTGATGAGCTCCGAGTCGATCATCCAGTGGCTGGAGCTGAGATCATCTCAGATG 500  
2Y 492 GCTGGTCTCTAAAGGAGCAACTACAGTGAATCTTAGCAAGATTTTAAAGAACTTTG 551

501 CTTGGTCTCTAAAGGAGCAACTACAGTGAATCTTAGCAAGATTTTAAAGAACTTTG 560  
552 ATAAATGAGTATTCGCGCTGGAATCCAGTGAATCTTGAAGTTTGGAAACAGATGAAA 611  
561 ATAAATGAGTATTCGCGCTGGAATCCAGTGAATCTTGAAGTTTGGAAACAGATGAAA 620  
612 CTGCAATCAATATTACAAAGTATCCAGACGAGGAGAAACAAGTTGCTTGAAGGAGCTCT 671  
621 CTGCAATCAATATTACAAAGTATCCAGACGAGGAGAAACAAGTTGCTTGAAGGAGCTCT 680  
672 GTGTGATGCTCTCAATTTGAATCTGATGAAAATGCTGATTTGAAAACCTCAGCTTCCAAG 731  
681 GTGTGATGCTCTCAATTTGAATCTGATGAAAATGCTGATTTGAAAACCTCAGCTTCCAAG 740  
732 AGTTTCTCAAGTGGCTCAACCCCATCTTTCAACCTCTGAGAAAGTGTGCCCTGGAGG 791  
741 AGTTTCTCAAGTGGCTCAACCCCATCTTTCAACCTCTGAGAAAGTGTGCCCTGGAGG 800  
792 ATGAAAAGTATGAGATGAGAGTGAAGCCGAGGTGAGATCTAACCCTGTGTGTGCTCT 851  
801 ATGAAAAGTATGAGATGAGAGTGAAGCCGAGGTGAGATCTAACCCTGTGTGTGCTCT 860  
852 GTGCAAAATGGGCTGTACAGCCATGAGCTGTGAGAAAGATCAGAAAGGGGCCCGAGA 911  
861 GTGCAAAATGGGCTGTACAGCCATGAGCTGTGAGAAAGATCAGAAAGGGGCCCGAGA 920  
912 CCCAGACAGAGGAGAGATGACAGATATGTCCAGGAGTCTCCAAAAGCATCAGGAAAACAG 971  
921 CCCAGACAGAGGAGAGATGACAGATATGTCCAGGAGTCTCCAAAAGCATCAGGAAAACAG 980  
972 CTGAAAAGCCAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 1031  
981 CTGAAAAGCCAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 1040  
1032 TGGATCCAGCATCTTCTCCATTCAGCGCTGAGTTCAGTATACAAAGTGTGTGTGCTCT 1091  
1041 TGGATCCAGCATCTTCTCCATTCAGCGCTGAGTTCAGTATACAAAGTGTGTGTGCTCT 1100  
1092 GTCCGCAATCAGAGTATTTGCTTATATAGCAATGAGTTTATTTTATTTTATTTT 1151  
1101 GTCCGCAATCAGAGTATTTGCTTATATAGCAATGAGTTTATTTTATTTTATTTT 1160  
1152 GCAATAAAGATATGAAGTGGCTGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1211  
1161 GCAATAAAGATATGAAGTGGCTGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1220  
1212 GAGTGTCTTAAGAGAAATCTGAAATGCTCTGCGGCTGAGGCTAGTAAAGAACTGTC 1271  
1221 GAGTGTCTTAAGAGAAATCTGAAATGCTCTGCGGCTGAGGCTAGTAAAGAACTGTC 1280  
1272 ATCAGATTTGAAAGAGAAACAGACCCCAATCTGAACTCTTTTGAAGTTTACTGCTATCTGT 1331  
1281 ATCAGATTTGAAAGAGAAACAGACCCCAATCTGAACTCTTTTGAAGTTTACTGCTATCTGT 1340  
1332 CAGCAGGCTGAGGAGTGCACACGATGCGAGAGAGAACTTAGCAGGCTGTCCCGGAGG 1391  
1341 CAGCAGGCTGAGGAGTGCACACGATGCGAGAGAGAACTTAGCAGGCTGTCCCGGAGG 1400  
1392 AGAGTTTGGGAAGCTCCAGGAGAGAAAGCTCTGCTTCCAGGCTCTTTTCCATTTGCC 1451  
1401 AGAGTTTGGGAAGCTCCAGGAGAGAAAGCTCTGCTTCCAGGCTCTTTTCCATTTGCC 1460  
1452 GTGAGATGACAGACCTCCAGCATCCAGATCCAGATCTCTTTGGTCCCAATTAAGTCCCTCTAGAT 1511  
1461 GTGAGATGACAGACCTCCAGCATCCAGCATCCAGATCTCTTTGGTCCCAATTAAGTCCCTCTAGAT 1520  
1512 ACATAGCCATCTAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1571  
1521 ACATAGCCATCTAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1580  
1572 GTACACCCAAATGATGAGATATTTGATATCTTTGAGCCCTTAGCCAGCTTAACCAATTT 1631



1581 GTACACCAAAATGATGAGATACCTGTATACCTTTAGAGCCCTTAGCAGCACTAACCAAAATT 1640  
1632 TTAATAAATACCTTTTACAAAGGTGCTATTCTCTGTAAACAACTTTTGTGGCAAGTT 1691  
1641 TTAATAAATACCTTTTACAAAGGTGCTATTCTCTGTAAACAACTTTTGTGGCAAGTT 1699  
1692 GACTTATTTCTCAAT 1751  
1700 GACTTATTTCTCAAT 1759  
1752 TAGTATTAAGCTTTTGTAAATATATTTTTCAGTAGTCCACCACTTCCATAGGTGGAGGAG 1811  
1760 TAGTATTAAGCTTTTGTAAATATATTTTTCAGTAGTCCACCACTTCCATAGGTGGAGGAG 1819  
1812 TTTGGGGTCTTCTCGTGTGAGGGGTGAAATAAACCAAGATGTCCTCCACCCCTGCCACATA 1871  
1820 TTTGGGGTCTTCTCGTGTGAGGGGTGAAATAAACCAAGATGTCCTCCACCCCTGCCACATA 1879  
1872 CTAGATGCAGCCCATAGTTGGCCCCCTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG 1931  
1880 CTAGATGCAGCCCATAGTTGGCCCCCTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG 1939  
1932 CAAGGGTGCTTAGACCTAAGCCAGGGGAAGAAGCATCTT 1971  
1940 CAAGGGTGCTTAGACCTAAGCCAGGGGAAGAAGCATCTT 1979

## RESULT 3

ABN96790 standard; DNA; 1987 BP.

ID ABN96790;

AC ABN96790;

CX 13-AUG-2002 (first entry)

CX Gene #3288 used to diagnose liver cancer.

CX Gene; liver cancer; db; hepatocellular carcinoma; hepatotropic;

CX metastatic liver tumour; cytostatic; expression profile; disease state;

CX disease progression; drug toxicity; drug efficacy; drug metabolism.

CX Homo sapiens.

CX MO20229103-A2.

CX 11-APR-2002.

CX 02-OCT-2001; 2001MO-US030589.

CX 02-OCT-2000; 2000US-0237054P.

CX (GENE-) GENE LOGIC INC.

CX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

CX WPI; 2002-426119/45.

CX Diagnosing and detecting the progression of liver cancer, hepatocellular

CX carcinoma or metastatic liver tumor in a patient, involves detecting the

CX level of expression of two or more genes in a liver tissue sample.

CX Claim 1; SEQ ID NO 3288; 298pp; English.

CX The invention relates to a novel method for diagnosing and detecting the

CX progression of liver cancer, hepatocellular carcinoma or metastatic liver

CX tumour in a patient, and differentiating metastatic liver cancer from

CX hepatocellular carcinoma in a patient, involving detecting the level of

CX expression of two or more genes represented in ABN93503-ABN97455 in a

CX tissue sample. The method of the invention has hepatotropic, and

CX cytostatic activity. The method is useful for diagnosing and detecting

CX the progression of liver cancer, hepatocellular carcinoma and metastatic

CX liver carcinoma in a patient. The method is useful for identifying

CX expression profiles which serve as useful diagnostic markers as well as

CC markers that can be used to monitor disease states, disease progression,  
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1987 BP; 522 A; 482 C; 498 G; 485 T; 0 U; 0 Other;

Query Match 52.5%; Score 1940; DB 6; Length 1987;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1954; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 12 CTCGCCACCTCGCTTACAGCTCGCTGCGCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGG 71  
DB 21 CTCGCCACCTCGCTTACAGCTCGCTGCGCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGG 80  
QY 72 ACCAGACACCATGTGAAACGCTGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCT 131  
DB 81 ACCAGACACCATGTGAAACGCTGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCT 140  
QY 132 GGGTCCGGCGCGAGAGAGCTAAGAGGCAATCCAGAGTCTGTGCCAATGTGTTTGTG 191  
DB 141 GGGTCCGGCGCGAGAGAGCTAAGAGGCAATCCAGAGTCTGTGCCAATGTGTTTGTG 200  
QY 192 GAGCGCGCGCGGAATGTGCAAGTCAAGAGAAAGGGAAACCCACCTGCTCTGCAATTGAGC 251  
DB 201 GAGCGCGCGCGGAATGTGCAAGTCAAGAGAAAGGGAAACCCACCTGCTCTGCAATTGAGC 260  
QY 252 AATGCAAACTCTCAAGAGGCTGTGTGCGAGTAATGGCAAGACCTACTCTCAACCACT 311  
DB 261 AATGCAAACTCTCAAGAGGCTGTGTGCGAGTAATGGCAAGACCTACTCTCAACCACT 320  
QY 312 GTGAACCTGCATCGAGATGCTGCTCACTGATCCAAATCCAGGTTGATTACGATGAC 371  
DB 321 GTGAACCTGCATCGAGATGCTGCTCACTGATCCAAATCCAGGTTGATTACGATGAC 380  
QY 372 ACTGCAAGAGAAATCCGTAACTCACTGCTCCAGCCCGAGTTGTTGCTATCAGTCCA 431  
DB 381 ACTGCAAGAGAAATCCGTAACTCACTGCTCCAGCCCGAGTTGTTGCTATCAGTCCA 440  
QY 432 ACCGTGATGAGCTCCGAGCTCGCATCATCCAGTGGCTGGAAGCTGAGATCATCTCAGATG 491  
DB 441 ACCGTGATGAGCTCCGAGCTCGCATCATCCAGTGGCTGGAAGCTGAGATCATCTCAGATG 500  
QY 492 GCTGTTCTCTTAAAGGCGAGCACTACAGTGAATCTCTAGACAAAGTATTTTAAAGAACTTTG 551  
DB 501 GCTGTTCTCTTAAAGGCGAGCACTACAGTGAATCTCTAGACAAAGTATTTTAAAGAACTTTG 560  
QY 552 ATAATGTTGATTTCTGCGCTGAGCTCCAGTGAATCTCTGAAAGTTTGTGGAACAGATGAAA 611  
DB 561 ATAATGTTGATTTCTGCGCTGAGCTCCAGTGAATCTCTGAAAGTTTGTGGAACAGATGAAA 620  
QY 612 CTGCCATCAATATTACAACTGATCCAGACCGAGGAGAAACAAAGTTGCTTAGGGGACTCT 671  
DB 621 CTGCCATCAATATTACAACTGATCCAGACCGAGGAGAAACAAAGTTGCTTAGGGGACTCT 680  
QY 672 GTGTTGATGCTCTCAATTGAACTGTGTGATGAAATGCTGATGAAACTCAGCTTCCAAAG 731  
DB 681 GTGTTGATGCTCTCAATTGAACTGTGTGATGAAATGCTGATGAAACTCAGCTTCCAAAG 740  
QY 732 AGTTTCTCAAGTGCCTCAACCCATCTTCAACCTCTGAGAGAAAGTGTGCCCTGGAGG 791  
DB 741 AGTTTCTCAAGTGCCTCAACCCATCTTCAACCTCTGAGAGAAAGTGTGCCCTGGAGG 800  
QY 792 ATGAAACGTTATGAGATGAGCTGAGACCGAGGTGAGCTGTAAACCGCTGTGTGTGCTGCT 851  
DB 801 ATGAAACGTTATGAGATGAGCTGAGACCGAGGTGAGCTGTAAACCGCTGTGTGTGCTGCT 860  
QY 852 GTGCAAAATTTGGTCTGTACAGCCATGCTGTGCGGAAAGAAATCAGAAAGGGGCCCGAGA 911  
DB 861 GTGCAAAATTTGGTCTGTACAGCCATGCTGTGCGGAAAGAAATCAGAAAGGGGCCCGAGA 920  
QY 912 CCCAGACAGAGGAGGAGATGACAGATATGTCCAGGAGCTTCCAAAGCATCAGGAAACAG 971

[illegible]

ABZ34719
ID ABZ34719 standard; cDNA; 2538 BP.
XX AC
XX AC ABZ34719;
XX DT
DT 04-FEB-2003 (first entry)
XX DE
DE Coding sequence SEQ ID 77, upregulated in osteogenesis.
XX KW
KW Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;
KW osteoporosis; bone disease; downregulator; human; ss.
XX OS
OS Homo sapiens.
XX FN
FN WO200281745-A2.
XX PD
PD 17-OCT-2002.
XX FF
FF 05-APR-2002; 2002WO-IB002211.
XX PR
PR 05-APR-2001; 2001US-0281400P.
XX PA
PA (AVET ) AVENTIS PHARMA SA.
XX PI
PI Garcia T. Roman Roman S, Baron R, Call K, Theilhaber J;
PI Connolly T, Jackson A, Bushnell SE, Rawadi G;
XX DR
DR WPI; 2003-058567/05.
XX PT
PT Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
PT useful for bone disease therapy in subject.
XX PS
PS Claim 26; Page 115-116; 237pp; English.
XX CC
CC The present invention relates to novel nucleotide sequences, which are
CC differentially expressed in models of osteogenesis upon being put in
CC contact with a stimulator of osteogenesis. The present sequence is one
CC such sequence. This sequence can be used for diagnosing osteoporosis/bone
CC disease in a patient, promoting osteogenesis and/or preventing
CC osteoporosis/bone disease. The present sequence encodes a secreted
CC protein
XX SQ
SQ Sequence 2538 BP; 707 A; 584 C; 595 G; 652 T; 0 U; 0 Other;
Query Match 46.8%; Score 1729.8; DB 7; Length 2538;
Best Local Similarity 99.8%; Pred.No. 0;
Matches 1742; Conservative 0; Mismatches 2; Indels 1; Gaps 1
QY 12 CTCGCCACTCGCTTACAGTCGCTGC CGCGCTCTGCCCGCGCCCCCAGAGACCTGG 71
Db 6 CTCGCCACTCGCTTACAGTCGCTGC CGCGCTCTGCCCGCGCCCCCAGAGACCTGG 65
QY 72 ACCAGACACGATGTGAACCGTGCTGCGCTCGCGCTCGCGCTGCGGTGGCGGTGCGCT 131
Db 66 ACCAGACACGATGTGAACCGTGCTGCGCTCGCGCTCGCGCTGCGGTGGCGGTGCGCT 125
QY 132 GGGTCCCGCGCGAGAGAGCTAAGGAGCAAATCCAAGATCTGTGCCAAATGTGTTTTGTG 191
Db 126 GGGTCCCGCGCGAGAGAGCTAAGGAGCAAATCCAAGATCTGTGCCAAATGTGTTTTGTG 185
QY 192 GAGCCGCGCGGGAATGTGCAGTCA CAGAGAAAGGGGAACCCACTGTCTCTGCATTGAGC 251
Db 186 GAGCCGCGCGGGAATGTGCAGTCA CAGAGAAAGGGGAACCCACTGTCTCTGCATTGAGC 245
QY 252 AATGCAAACTCACAGAGGCGCTGTGTGTGGCAGTAATGGCAAGACCTTACCTCAACCACT 311
Db 246 AATGCAAACTCACAGAGGCGCTGTGTGTGGCAGTAATGGCAAGACCTTACCTCAACCACT 305
QY 312 GTGAATGCGATCGAGATGCGCTGCTCATCGGATCCAAAATCCAGGTGATTCAGATGGAC 371
Db 306 GTGAATGCGATCGAGATGCGCTGCTCATCGGATCCAAAATCCAGGTGATTCAGATGGAC 365
QY 372 ACTGCAAGAGAGAAGAAATCCGTTAAGTCCATCTGCCAGCCCAAGTTGTTTGTCTATCAGTCCA 431

366 ACTGCAAGAGAGAAATCCGTAAGTCCATCTGCCAGCCAGTGTGTGCTATCAGTCCA 425  
 432 ACCGTGATGAGTCCGACGTCGATCATCATCAGTGGCTGGAGCTGAGATCATCTCAGATG 491  
 426 ACCGTGATGAGTCCGACGTCGATCATCATCAGTGGCTGGAGCTGAGATCATCTCAGATG 485  
 492 GCTGGTCTCTTAAGGAGGAGCAACTACAGTGAATCTTAGCAAGATATTTAAGAACTTTG 551  
 486 GCTGGTCTCTTAAGGAGGAGCAACTACAGTGAATCTTAGCAAGATATTTAAGAACTTTG 545  
 552 ATATGCTGATCTCTGCGCTGGACTCCAGTCAATCTCTGAACTTTTGGACACAAATGAAA 611  
 546 ATATGCTGATCTCTGCGCTGGACTCCAGTGAATCTCTGAACTTTTGGACACAAATGAAA 605  
 612 CTGCAATCAATATTAACAGTATCCAGACAGAGAGAAACAAGTGTCTTAGGGACTCT 671  
 606 CTGCAATCAATATTAACAGTATCCAGACAGAGAGAAACAAGTGTCTTAGGGACTCT 665  
 672 GTGTTGATGCTCTCATTAAGTCTGTGATGAATCTGATGGAATCTGATGGAATCTGATG 731  
 666 GTGTTGATGCTCTCATTAAGTCTGTGATGAATCTGATGGAATCTGATGGAATCTGATG 725  
 732 AGTTTCTCAAGTGCCTCAACCCCTCTTTCAACCCCTCTGAGAGAAAGTGTGCCCTGGAG 791  
 726 AGTTTCTCAAGTGCCTCAACCCCTCTTTCAACCCCTCTGAGAGAAAGTGTGCCCTGGAG 785  
 792 ATGAAAGTATGAGATGGAGCTGAGACCGAGTGGACTGTACCGCTGTGTGTGCT 851  
 786 ATGAAAGTATGAGATGGAGCTGAGACCGAGTGGACTGTACCGCTGTGTGTGCT 845  
 852 GTGAAATTTGGTCTGTACAGCATGACCTGTGACCGAAAGATCAAGAGGGGGCCAG 911  
 846 GTGAAATTTGGTCTGTACAGCATGACCTGTGACCGAAAGATCAAGAGGGGGCCAG 905  
 912 CCAGACAGAGAGGAGATACAGAGATATGTCCAGAGCTCCAAAGCATCAGGAACAG 971  
 906 CCAGACAGAGAGGAGATACAGAGATATGTCCAGAGCTCCAAAGCATCAGGAACAG 965  
 972 CTGAAAGACACAGAGAGTGAAGACCAAGAGATCTAATGAGGAGGACAGACAGCTGTC 1031  
 966 CTGAAAGACACAGAGAGTGAAGACCAAGAGATCTAATGAGGAGGACAGACAGCTGTC 1025  
 1032 TGGATCCAGATCTTCTCCACTTCAGCTGAGTTCAGTATACAGAGTGTCTGTACA 1091  
 1026 TGGATCCAGATCTTCTCCACTTCAGCTGAGTTCAGTATACAGAGTGTCTGTACA 1085  
 1092 GTGCCCAATCACCAGTATTTGCTTATATAGCAATGATTTTATTTTGTCTTTT 1151  
 1086 GTGCCCAATCACCAGTATTTGCTTATATAGCAATGATTTTATTTTGTCTTTT 1145  
 1152 GCAATAAAGGATATGAAGTGGCTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1211  
 1146 GCAATAAAGGATATGAAGTGGCTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1205  
 1212 GAGTGTCTTTAAGAGAACTGTAAATGCTGCTGAGGCTGAGGCTAGTAAGGAGAACTGC 1271  
 1206 GAGTGTCTTTAAGAGAACTGTAAATGCTGCTGAGGCTGAGGCTAGTAAGGAGAACTGC 1265  
 1272 ATCAGGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1331  
 1266 ATCAGGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1325  
 1332 CAGCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1391  
 1326 CAGCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1385  
 1392 AGAGGTTTGGAGAGCTCCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1451  
 1386 AGAGGTTTGGAGAGCTCCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1445  
 1452 GTCAGCATGACAGACCTCCAGCATCCAGCATCTCTTGGTCCCAATTAATGCTCCTAGAT 1511

1446 GTCAGCATGACAGACCTCCAGCATCCAGCATCTCTGGTCCCAATTAATGCTCCTAGAT 1505  
 1512 ACATAGCCATATCTGCTAGTTAAACCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGGAGG 1571  
 1506 ACATAGCCATATCTGCTAGTTAAACCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGGAGG 1565  
 1572 GTACACCCAAATGATGAGATCTGTATATCTTTGAGCCCTTAGCGACCTTAACCAATTT 1631  
 1566 GTACACCCAAATGATGAGATCTGTATATCTTTGAGCCCTTAGCGACCTTAACCAATTT 1625  
 1632 TTAAAAATACCTTTTACCAGAGTGTCTATTTCTCTGTAAACAC-TTTTTTTGGCAAGT 1690  
 1626 TTAAAAATACCTTTTACCAGAGTGTCTATTTCTCTGTAAACACCTTTTTTTTGGCAAGT 1685  
 1691 TGACTTTATCTTCAATATTTATATATATATATATATATATATATATATATATATATTTTCTTGA 1750  
 1686 TGACTTTATCTTCAATATTTATATATATATATATATATATATATATATATATATTTTCTTGA 1745  
 1751 CTAGG 1755  
 1746 CTAGG 1750

RESULT 5  
 AAI57940/c  
 ID AAI57940 standard; cDNA; 2655 BP.  
 XX  
 AC AAI57940;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 143.  
 XX  
 KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US034263.  
 XX  
 PR 23-DEC-1999; 98US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00598042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX  
 (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 PI Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR P-PSDB; AAM38784.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.  
 XX  
 PS Claim 1; SEQ ID NO 143; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57998-AAI61369) and the





05-DEC-2000; 2000US-0251030P.  
 05-DEC-2000; 2000US-0251988P.  
 05-DEC-2000; 2000US-0256719P.  
 06-DEC-2000; 2000US-0251479P.  
 08-DEC-2000; 2000US-0251856P.  
 08-DEC-2000; 2000US-0251868P.  
 08-DEC-2000; 2000US-0251889P.  
 08-DEC-2000; 2000US-0251989P.  
 08-DEC-2000; 2000US-0251990P.  
 11-DEC-2000; 2000US-0254097P.  
 05-JAN-2001; 2001US-0259678P.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Rosen CA, Barash SC, Ruben SM;  
 WPI; 2001-451937/48.  
 Isolated polypeptide for treating, preventing and/or prognosing  
 disorders related to the musculoskeletal system including musculoskeletal  
 cancers and also for testing and detection e.g. diagnosis.  
 Example 2; SEQ ID NO 3903; 781pp + Sequence Listing; English.  
 The invention relates to novel genes (AAL34659-AAL37666) and proteins  
 (ABB03087-ABB04109) associated with the musculoskeletal system useful for  
 preventing, treating or ameliorating medical conditions e.g. by protein  
 or gene therapy. The genes are isolated from a range of human tissues  
 disclosed in the specification. The nucleic acids, proteins, antibodies  
 and (ant)agonists are useful in the diagnosis, treatment and prevention  
 of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the  
 adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
 lung, or urogenital; (b) immune disorders e.g. Addison's disease,  
 allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 arthritis and ulcerative colitis; (c) cardiovascular disorders such as  
 myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.  
 cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,  
 bacterial, fungal and parasitic infections. Note: The sequence data for  
 this patent did not form part of the printed specification, but was  
 obtained in electronic format directly from WIPO at  
 ftp.wipo.int/pub/published\_pct\_sequences  
 Sequence 1755 BP; 500 A; 356 G; 351 G; 546 T; 0 U; 2 Other;

Query Match 45.4%; Score 1678.4; DB 4; Length 1755;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 1682; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 1993 AACACATTAATTTGTTTATTTATTTCTGAGAGTTGAGGCAATTCAGTATCCCAAGGAT 2052  
 1 ANANTTTTGTGTTTTTATTTATTTCTGAGAGTTGAGGCAATTCAGTATCCCAAGGAT 60  
 2053 GCGCAAGGGCAGCAAGCAGGGCTTAGGATATCCAGCCTACCAATATGCTCATTCGA 2112  
 61 GCGCAAGGGCAGCAAGCAGGGCTTAGGATATCCAGCCTACCAATATGCTCATTCGA 120  
 2113 CTAACATAGAGGGTGAAGTGGCCCTGCTCTCTTTTCTGGACCTCAGTTCTTCAGT 2172  
 121 CTAACATAGAGGGTGAAGTGGCCCTGCTCTCTTTTCTGGACCTCAGTTCTTCAGT 180  
 2173 GAGCTGGTAAGAAATGCACATACCTTTGATTGTAAGTATATAAATCTGTGGTTCTGAT 2232  
 181 GAGCTGGTAAGAAATGCACATACCTTTGATTGTAAGTATATAAATCTGTGGTTCTGAT 240  
 2233 CATTGGTCCAGAGGGAGATAGTTCCTGATTTTCTTTCTCTATAGATAAATG 2292  
 241 CATTGGTCCAGAGGGAGATAGTTCCTGATTTTCTTTCTCTATAGATAAATG 300  
 2293 AAATCTTGTTACTAGAACAAAGAAATGTCAGATGGCCAAAAACAAGATGACCAAGTTTGTAT 2352  
 301 AAATCTTGTTACTAGAACAAAGAAATGTCAGATGGCCAAAAACAAGATGACCAAGTTTGTAT 360  
 2353 CTCAGCCTGATGACCCCTACAGGTCGTGCTATGATATGGAGTCCTCATGGGTAAGCAGGA 2412  
 361 CTCAGCCTGATGACCCCTACAGGTCGTGCTATGATATGGAGTCCTCATGGGTAAGCAGGA 420  
 2413 AGAGAGTGGGAAAGAGAACACCCACCTCTGCTTCATATTTGATTTTCATGTTTAACT 2472  
 421 AGAGAGTGGGAAAGAGAACACCCACCTCTGCTTCATATTTGATTTTCATGTTTAACT 480  
 2473 CCGGCTGGAATAGAAACATTCCTTTAGAGATGAGGATAAAAGAAAGTTTCAGATTCAA 2532  
 481 CCGGCTGGAATAGAAACATTCCTTTAGAGATGAGGATAAAAGAAAGTTTCAGATTCAA 540  
 2533 CAGGGGGAAGAAATGGAGATTTAATCCTAAAACTGTGACTTGGGAGGTCAGTCATTTA 2592  
 541 CAGGGGGAAGAAATGGAGATTTAATCCTAAAACTGTGACTTGGGAGGTCAGTCATTTA 600  
 2593 CAGTTAGTCCTGTGCTTTTCGACTTCTGTGATTTAATACCCACTCACTACCCCTTTCA 2652  
 601 CAGTTAGTCCTGTGCTTTTCGACTTCTGTGATTTAATACCCACTCACTACCCCTTTCA 660  
 2653 GATGCATTTGGAAATACCAAGATTAATCCTTTGACATAGATCTCATTTGAGAAAGCAG 2712  
 661 GATGCATTTGGAAATACCAAGATTAATCCTTTGACATAGATCTCATTTGAGAAAGCAG 720  
 2713 ATTAAGACCATCAGAGGAAATTTTAGTGTGTAATGCAAGCAACTGTGAGAAACT 2772  
 721 ATTAAGACCATCAGAGGAAATTTTAGTGTGTAATGCAAGCAACTGTGAGAAACT 780  
 2773 GTGTGCCCAAAATAGAAATTCCTTCTAGTCTTTCTGTTCTCATTTGAAAGAGGAAAT 2832  
 781 GTGTGCCCAAAATAGAAATTCCTTCTAGTCTTTCTGTTCTCATTTGAAAGAGGAAAT 840  
 2833 CCACCTTTGTTAGCAATTTCAAGCTTTTATGATCCATCCCATCTAAAAAATCTTCAAACT 2892  
 841 CCACCTTTGTTAGCAATTTCAAGCTTTTATGATCCATCCCATCTAAAAAATCTTCAAACT 900  
 2893 CCACCTTTGTTAGCAATTTCAAGCTTTCTGTCAGTCCCTGCAAGTCCCTGGAGAACTCAGAGCAGC 2952  
 901 CCACCTTTGTTAGCAATTTCAAGCTTTCTGTCAGTCCCTGTCAGTCCCTGGAGAACTCAGAGCAGC 960  
 2953 AGCCCTTAATCAAGGTTTACAGCCCTTTGAGCACTATGGAGGAGGAGGCAAGAGTACAC 3012  
 961 AGCCCTTAATCAAGGTTTACAGCCCTTTGAGCACTATGGAGGAGGAGGCAAGAGTACAC 1020  
 3013 CAATTTGTTAAAGCAAGAAACCAAGTCTCTTCTACATAGTCAATTTAGAACATGTTTAT 3072  
 1021 CAATTTGTTAAAGCAAGAAACCAAGTCTCTTCTACATAGTCAATTTAGAACATGTTTAT 1080  
 3073 CATCAAGACTACTCTACCCCTGCAACATTTGAATCCCAAGAGCAATCCACATTCCTCTT 3132  
 1081 CATCAAGACTACTCTACCCCTGCAACATTTGAATCCCAAGAGCAATCCACATTCCTCTT 1140  
 3133 GAGTTCTGCAGCTTCTGTGTAATAGGCGAGTGTGCTATGCGGTAGAAATCAGATGAT 3192  
 1141 GAGTTCTGCAGCTTCTGTGTAATAGGCGAGTGTGCTATGCGGTAGAAATCAGATGAT 1200  
 3193 CTGAGGACCATTCATGGAAGTCTGCTAAATAGCTAGTCTGGGGAGTCTTCCATAAGTTT 3252  
 1201 CTGAGGACCATTCATGGAAGTCTGCTAAATAGCTAGTCTGGGGAGTCTTCCATAAGTTT 1260  
 3253 TGCATGGAGCAAAACAAACAGGATTAATAGTGTGTTTCTTCAGCCCTCTAAAGCAT 3312  
 1261 TGCATGGAGCAAAACAAACAGGATTAATAGTGTGTTTCTTCAGCCCTCTAAAGCAT 1320  
 3313 AGGGCTTAGCCCTGAGGCTTCTTGGGCTTCTCTGCTGTGTAGTTTGTGTAACACATAT 3372  
 1321 AGGGCTTAGCCCTGAGGCTTCTTGGGCTTCTCTGCTGTGTAGTTTGTGTAACACATAT 1380  
 3373 AGCATCTGTTAAGATCCAGTGTCCAATGGAAACCTTCCACATGCGGTGACTCTGGACTAT 3432  
 1381 AGCATCTGTTAAGATCCAGTGTCCAATGGAAACATTCCACATGCGGTGACTCTGGACTAT 1440  
 3433 ATCAGTTTTTGGAAAGCAGGGTTCTCTGCTGCTGTCTAAACAGCCACGTGGACCATCTGA 3492



1441 ATGATTTTGGAAAGCAGGGTCCCTGCTGCTACACAGCCAGTGACAGTCTGA 1500  
3493 ATGCTTTTCTTACACTAGTATTTTAACTAGTCAAACTTCAAGAAACAATCTAAACAA 3552  
1501 ATGCTTTTCTTACACTAGTATTTTAACTAGTCAAACTTCAAGAAACAATCTAAACAA 1560  
3553 GTTCTCTGTCATATGTTTGGAACTTCTATTTTATTTAGTCTTCTATATTGCA 3612  
1561 GTTCTCTGTCATATGTTTGGAACTTCTATTTTATTTAGTCTTCTATATTGCA 1620  
3613 TTTAACTTGTTTTGTAACTTCTGATTTCTTCTTCTTGGTACTATTTGATGAATAAGAA 3672  
1621 TTTAACTTGTTTTGTAACTTCTGATTTCTTCTTCTTGGTACTATTTGATGAATAAGAA 1680  
3673 ATTAAGTGA 3682  
1681 ATTAAGTGA 1690

RESULT 7  
ABX60526  
ID ABX60526 standard; cDNA; 1755 BP.  
CX  
CX ABX60526;  
CX  
CX  
CX 26-FEB-2003 (first entry)  
CX cDNA encoding novel human musculoskeletal system antigen #2870.  
CX Gene; ss; musculoskeletal system antigen; cancer; metastasis;  
CX re-vascularisation; thrombosis; arteriosclerosis; mineral content;  
CX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;  
CX post-operative tissue repair; limb regeneration; neuronal growth;  
CX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
CX AIDS-related complex; chondrocyte growth; bone regeneration;  
CX periodontal regeneration; tissue transport; bone graft; skin aging;  
CX keratinocyte growth; hair loss; melanocyte growth; cell proliferation;  
CX cell growth; organ transplant; cell differentiation; body height; weight;  
CX hair colour; eye colour; skin; percentage of adipose tissue;  
CX pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;  
CX depression; tendency for violence; pain; reproductive capability;  
CX hormone level; endocrine level; appetite; libido; memory; stress;  
CX storage capability; fat content; lipid content; protein content;  
CX carbohydrate content; vitamin content; cofactor content;  
CX nutritional component.  
CX  
CX Homo sapiens.  
CX  
CX US2002147140-A1.  
CX  
CX 10-OCT-2002.  
CX  
CX 17-JAN-2001; 2001US-00764877.  
CX  
CX 31-JAN-2000; 2000US-0179065P.  
CX 04-FEB-2000; 2000US-0180628P.  
CX 28-JUN-2000; 2000US-0214896P.  
CX 07-JUL-2000; 2000US-0216647P.  
CX 07-JUL-2000; 2000US-0216880P.  
CX 11-JUL-2000; 2000US-0217487P.  
CX 11-JUL-2000; 2000US-0217496P.  
CX 14-JUL-2000; 2000US-0218230P.  
CX 26-JUL-2000; 2000US-0220963P.  
CX 26-JUL-2000; 2000US-0220964P.  
CX 14-AUG-2000; 2000US-0224518P.  
CX 14-AUG-2000; 2000US-0224519P.  
CX 14-AUG-2000; 2000US-0225267P.  
CX 14-AUG-2000; 2000US-0225268P.  
CX 14-AUG-2000; 2000US-0225270P.  
CX 14-AUG-2000; 2000US-0225447P.  
CX 14-AUG-2000; 2000US-0225757P.  
CX 14-AUG-2000; 2000US-0225758P.

22-AUG-2000; 2000US-0226868P.  
30-AUG-2000; 2000US-0228924P.  
01-SEP-2000; 2000US-0229287P.  
01-SEP-2000; 2000US-0229343P.  
01-SEP-2000; 2000US-0229344P.  
01-SEP-2000; 2000US-0229345P.  
05-SEP-2000; 2000US-0229509P.  
05-SEP-2000; 2000US-0229513P.  
08-SEP-2000; 2000US-0231413P.  
21-SEP-2000; 2000US-0234233P.  
21-SEP-2000; 2000US-0234274P.  
25-SEP-2000; 2000US-0234997P.  
27-SEP-2000; 2000US-0235834P.  
29-SEP-2000; 2000US-0236327P.  
29-SEP-2000; 2000US-0236367P.  
29-SEP-2000; 2000US-0236382P.  
29-SEP-2000; 2000US-0236389P.  
29-SEP-2000; 2000US-0236370P.  
02-OCT-2000; 2000US-0236802P.  
02-OCT-2000; 2000US-0237037P.  
02-OCT-2000; 2000US-0237038P.  
02-OCT-2000; 2000US-0237039P.  
02-OCT-2000; 2000US-0237040P.  
13-OCT-2000; 2000US-0239352P.  
20-OCT-2000; 2000US-0240960P.  
20-OCT-2000; 2000US-0241785P.  
20-OCT-2000; 2000US-0241809P.  
01-NOV-2000; 2000US-0244617P.  
17-NOV-2000; 2000US-0249299P.  
08-DEC-2000; 2000US-0251856P.  
08-DEC-2000; 2000US-0251868P.  
08-DEC-2000; 2000US-0251869P.  
(ROSE/) ROSEN C A.  
(RUBE/) RUBEN S M.  
(BARA/) BARASH S C.  
Rosen CA, Ruben SM, Barash SC;  
WFI; 2003-128199/12.  
Isolated nucleic acid molecules encoding musculoskeletal system associated polypeptides, useful for detecting disorders, e.g. cancer.  
Disclosure; SEQ ID NO 3903; 321pp; English.  
The invention describes an isolated nucleic acid molecule comprising a sequence encoding musculoskeletal system associated polypeptides useful for detecting disorders, e.g., cancer or cancer metastases, in animals or humans. The nucleic acid: stimulates re-vascularisation of ischaemic tissues associated with conditions such as thrombosis, arteriosclerosis, and other cardiovascular conditions; treats wounds due to injuries, burns, post-operative tissue repair, and ulcers; stimulates angiogenesis and limb regeneration; stimulates neuronal growth; can treat and prevent neuronal damage occurring in certain disorders or neurodegenerative conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-related complex; stimulates chondrocyte growth, thus they can be used to enhance bone and periodontal regeneration and aid in tissue transports or bone grafts; prevents skin aging due to sunburn by stimulating keratinocyte growth; prevents hair loss, since FGF family members activate hair-forming cells and promotes melanocyte growth; stimulates growth and differentiation of hematopoietic cells and bone marrow cells when used in combination with other cytokines; maintains organs before transplantation or for supporting cell culture of primary tissues; induces tissue of mesodermal origin to differentiate in early embryos; increases or decreases the differentiation or proliferation of embryonic stem cells, besides, hematopoietic lineage; modulates mammalian characteristics, such as, body height, weight, hair colour, eye colour, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery); modulates mammalian metabolism; changes mammal's metal state or physical state by influencing biorhythms, circadian rhythms, depression, tendency for violence, tolerance for pain, reproductive capabilities, hormonal or endocrine levels, appetite, libido, memory, or

CC stress; increases or decreases storage capabilities, fat content, lipid,  
CC protein, carbohydrate, vitamins, minerals, cofactors or other nutritional  
CC components. This sequence encodes a novel human musculoskeletal system  
CC antigen. Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from the US patent office at  
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140  
XX  
XX  
SQ Sequence 1755 BP; 500 A; 356 C; 351 G; 546 T; 0 U; 2 Other;

Query Match 45.4%; Score 1678.4; DB 7; Length 1755;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1682; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

2y 1993 AACATTAATTTGTTTATTTATTTCTGAGAAGTTGAGGCAATTCAGTATTCCTCAAGGAT 2052  
Db 1 ANANNTTTTGTGTTTTTATTTATTTCTGAGAAGTTGAGGCAATTCAGTATTCCTCAAGGAT 60

2y 2053 GGCACAAAGGCGACCAAGCAGGCTTAGGATATCCAGCCTACCAATATGCTCATTCGA 2112  
Db 61 GGCACAAAGGCGACCAAGCAGGCTTAGGATATCCAGCCTACCAATATGCTCATTCGA 120

2y 2113 CTAAGTGGAGGGTGAAGTGGCCCTGCTCTCTTTTCTGAGCCTCAAGTTCCTCAGT 2172  
Db 121 CTAAGTGGAGGGTGAAGTGGCCCTGCTCTCTTTTCTGAGCCTCAAGTTCCTCAGT 180

2y 2173 GAGCTGTAAGAATGCACTAACTTTTGAATTTGATAGTTTATAAATCTGTGTTCTGAT 2232  
Db 181 GAGCTGTAAGAATGCACTAACTTTTGAATTTGATAGTTTATAAATCTGTGTTCTGAT 240

2y 2233 CATTTGTCGAGGGGAGATAGTTTCTGTGATTTTCTTCTCTATAGATTAAGT 2292  
Db 241 CATTTGTCGAGGGGAGATAGTTTCTGTGATTTTCTTCTCTATAGATTAAGT 300

2y 2293 AAATCTTTGTTACTAGAACAAAGAAATGTCAGATGGCCAAAGAACAGATGATTTGAT 2352  
Db 301 AAATCTTTGTTACTAGAACAAAGAAATGTCAGATGGCCAAAGAACAGATGATTTGAT 360

2y 2353 CTCAGCTGATGACCTACAGGTCGCTATGATAGGATCTCTATGAGTAAAGCAGGA 2412  
Db 361 CTCAGCTGATGACCTACAGGTCGCTATGATAGGATCTCTATGAGTAAAGCAGGA 420

2y 2413 AGAGAGTGGGAAGAGAACCAACCCACTCTGTCTCATATTTGCAATTCATGTTAACT 2472  
Db 421 AGAGAGTGGGAAGAGAACCAACCCACTCTGTCTCATATTTGCAATTCATGTTAACT 480

2y 2473 CCGCTCGAAATAGAAAGCATTCCTTAGAGATGAGGATAAAGAAAGTTTCAGATTCAA 2532  
Db 481 CCGCTCGAAATAGAAAGCATTCCTTAGAGATGAGGATAAAGAAAGTTTCAGATTCAA 540

2y 2533 CAGGGGAAGAAATGAGATTTAATCCTAAAGTGTGACTTGGGAGGTCAGTCATTTA 2592  
Db 541 CAGGGGAAGAAATGAGATTTAATCCTAAAGTGTGACTTGGGAGGTCAGTCATTTA 600

2y 2593 CAGTTAGTCTGTCTTTGCACTTCTGTGATTTAAACCCCACTCACTACCTGTTTCA 2652  
Db 601 CAGTTAGTCTGTCTTTGCACTTCTGTGATTTAAACCCCACTCACTACCTGTTTCA 660

2y 2653 GATCATTTGGAATACCAAGATTTAAATCCTTGACATPAAGATCTATTTGCAAGAGCAG 2712  
Db 661 GATCATTTGGAATACCAAGATTTAAATCCTTGACATPAAGATCTATTTGCAAGAGCAG 720

2y 2713 ATTAAGACCATCAGAGGAAATTTATTAGTGTGTAATGACAGGCAACTGTGAGAACT 2772  
Db 721 ATTAAGACCATCAGAGGAAATTTATTAGTGTGTAATGACAGGCAACTGTGAGAACT 780

2y 2773 GTTCGCGCAAAATAGATTCCTCTAGTTTCTTCTGTTCTCTATTTGAAAGGAGAAAT 2832  
Db 781 GTTCGCGCAAAATAGATTCCTCTAGTTTCTTCTGTTCTCTATTTGAAAGGAGAAAT 840

2y 2833 CCACCTTTGTTAGCATTCAGCTTTTATGATCCATCCCATCTAAAGAACTCTTCAACT 2892  
Db 841 CCACCTTTGTTAGCATTCAGCTTTTATGATCCATCCCATCTAAAGAACTCTTCAACT 900

QY 2893 CCACCTTTGTTAGTCTGAAATGCAAGCTCCTCTGCAAGTGCCTTGGAGAACTCAGCAGC 2952  
Db 901 CCACCTTTGTTAGTCTGAAATGCAAGCTCCTCTGCAAGTGCCTTGGAGAACTCAGCAGC 960

QY 2953 AGCCCTTAATCAAAGGTTTATACCAAGCCCTTGGACACTATGGAGAGGAGGCAAGAGTACAC 3012  
Db 961 AGCCCTTAATCAAAGGTTTATACCAAGCCCTTGGACACTATGGAGAGGAGGCAAGAGTACAC 1020

QY 3013 CAATTTGTTAAAGCAAGAAACCAAGTGTCTCTCTACTAGTCAATTTAGAACATGGTTAT 3072  
Db 1021 CAATTTGTTAAAGCAAGAAACCAAGTGTCTCTCTACTAGTCAATTTAGAACATGGTTAT 1080

QY 3073 CATCCAGACTACTCTACCTGCAACATTGAATCCCAAGAGCAAAATCCACATTTCTCTCT 3132  
Db 1081 CATCCAGACTACTCTACCTGCAACATTGAATCCCAAGAGCAAAATCCACATTTCTCTCT 1140

QY 3133 GAGTCTCGAGCTTCTGTGTAATAGGAGCAGCTGTCTGTATGCGGTAGATACATGAT 3192  
Db 1141 GAGTCTCGAGCTTCTGTGTAATAGGAGCAGCTGTCTGTATGCGGTAGATACATGAT 1200

QY 3193 CTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGTCTGGGAGTCTTCCATAAAGTTT 3252  
Db 1201 CTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGTCTGGGAGTCTTCCATAAAGTTT 1260

QY 3253 TGCATGGAGCAAAACAAACAGGATTAACCTAGTTTGGTTCTTCAGCCCTCTAAAGCAT 3312  
Db 1261 TGCATGGAGCAAAACAAACAGGATTAACCTAGTTTGGTTCTTCAGCCCTCTAAAGCAT 1320

QY 3313 AGGGCTTAGCCTGAGGCTTCTTGGGCTTTCTGTGTGTGTAGTTTGTGTAACACTAT 3372  
Db 1321 AGGGCTTAGCCTGAGGCTTCTTGGGCTTTCTGTGTGTGTAGTTTGTGTAACACTAT 1380

QY 3373 AGCATCTCTTAAAGTCCAGTGTCCATGGAACCTTCCACATGCGGTGACTCTCGACTAT 3432  
Db 1381 AGCATCTCTTAAAGTCCAGTGTCCATGGAACCTTCCACATGCGGTGACTCTCGACTAT 1440

QY 3433 ATCAGTCTTTCGAAGCAGGCTTCTCTGCTGCTAAACAGCCCAAGTGGACCACTGTA 3492  
Db 1441 ATCAGTCTTTCGAAGCAGGCTTCTCTGCTGCTAAACAGCCCAAGTGGACCACTGTA 1500

QY 3493 ATGCTCTTCTTTACACCTATGTTTAAAGTAGTCAAACTTCAAGAAACAATCTAAACAA 3552  
Db 1501 ATGCTCTTCTTTACACCTATGTTTAAAGTAGTCAAACTTCAAGAAACAATCTAAACAA 1560

QY 3553 GTTCTGTGCTATGTTGTGAACCTGTTATTTAGTATTTAGTAGGCTTCTATATTGCA 3612  
Db 1561 GTTCTGTGCTATGTTGTGAACCTGTTATTTAGTATTTAGTAGGCTTCTATATTGCA 1620

QY 3613 TTTAACTTTGTTTGTAACTCCTGATTCCTCTTTCGGATCTATTGATGAATAAAGAA 3672  
Db 1621 TTTAACTTTGTTTGTAACTCCTGATTCCTCTTTCGGATCTATTGATGAATAAAGAA 1680

QY 3673 ATTAAGTGA 3682  
Db 1681 ATTAAGTGA 1690

RESULT 8  
ABQ88210  
ID ABQ88210 standard; cDNA; 1010 BP.  
XX  
AC ABQ88210;  
XX  
DT 18-SEP-2002 (first entry)  
XX  
XX Human osteoblast differentiation related cDNA SEQ ID NO 117.  
DE  
XX Human; osteoblast; stem cell differentiation; bone tissue deposition;  
KW osteoporosis; osteopathic; ss.  
XX  
OS Homo sapiens.  
XX

PN WO200250301-A2.  
 PD 27-JUN-2002.  
 PF 18-DEC-2001; 2001WO-US048276.  
 PR 18-DEC-2000; 2000US-0255882P.  
 PR 24-APR-2001; 2001US-0285691P.  
 XX (GENE-) GENE LOGIC INC.  
 PA (PROC) PROCTER & GAMBLE CO.  
 XX Ji D, Axelrod DW, Cook JS, Jaisswal N, Binstein R, Houghton A;  
 PI Mertz L;  
 XX WPI; 2002-557663/59.  
 DR Use of genes and their expression profiles associated with osteoblast  
 PT differentiation for screening modulators bone formation, for diagnosing  
 PT or treating e.g. osteoporosis, or as markers for the differentiation  
 PT process.  
 XX Claim 1; SEQ ID NO 117; 78pp + Sequence Listing; English.  
 PS The invention relates to genes and their expression profiles are used  
 CC for: (a) screening modulators of precursor stem cell differentiation into  
 CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal  
 CC deposition of bone tissue, abnormal rate of osteoblast formation or  
 CC osteoporosis; or (c) treating or monitoring treatment of the conditions  
 CC cited in (b), or monitoring the progression of bone tissue deposition.  
 CC Specific conditions include postmenopausal osteoporosis, glucocorticoid  
 CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-  
 CC induced abnormalities in bone formation or bone loss, conditions that  
 CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),  
 CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome  
 CC or fibrous dysplasia. The present sequence is that of an osteoblast  
 CC differentiation associated cDNA marker of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 1010 BP; 295 A; 211 C; 190 G; 314 T; 0 U; 0 Other;  
 SQ

Query Match 27.3%; Score 1008.4; DB 6; Length 1010;  
 Best Local Similarity 99.9%; Pred. No. 2.2e-248;  
 Matches 1009; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 2683 TTGACATAGATCTCATTTGCGAAGAGCAGATTAAAGACCATCAGAGGAAATTTATTAG 2742  
 1 TTGACATAGATCTCATTTGCGAAGAGCAGATTAAAGACCATCAGAGGAAATTTATTAG 60  
 2743 GTTGTAAATGCGACGCAACTGTGAAACTTTGTGCGAAGAGCAGATTAAAGACCATCAGAGGAAATTTATTAG 2802  
 61 GTTGTAAATGCGACGCAACTGTGAAACTTTGTGCGAAGAGCAGATTAAAGACCATCAGAGGAAATTTATTAG 120  
 2803 TTTCTTGTCTCATTTGAAAGAGGAAATTCACATTTGTTTGTAGCATTTCAAGCTTTATG 2862  
 121 TTTCTTGTCTCATTTGAAAGAGGAAATTCACATTTGTTTGTAGCATTTCAAGCTTTATG 180  
 2863 TATCCATCCCATCTATAAAGCTTTCAAACTCCACTTTGTTGAGTCTGAAATGCGAGTCCCT 2922  
 181 TATCCATCCCATCTATAAAGCTTTCAAACTCCACTTTGTTGAGTCTGAAATGCGAGTCCCT 240  
 2923 GTCCAGTGCCTTGGAGCACTCAGAGCAGCGCTTAATCAAGGTTTATACGAGCCCTT 2982  
 241 GTCCAGTGCCTTGGAGCACTCAGAGCAGCGCTTAATCAAGGTTTATACGAGCCCTT 300  
 2983 GGACACTATGGGAGGAGGAGGAGTACACCAATTTGTTAAAGCAAGAAACACACAGTGT 3042  
 301 GGACACTATGGGAGGAGGAGGAGTACACCAATTTGTTAAAGCAAGAAACACACAGTGT 360  
 3043 CTCCTCAGTATGATTTAGAACATGGTTATCATCCAGCACTCTACCTCGACATTTG 3102

Db 361 CTCCTCAGTATGATTTAGAACATGGTTATCATCCAGACTACTCTACCTGCAACATTTG 420  
 QY 3103 AACTCCCAAGAGCAAAATCCACATTCCTCTTGAGTTCTCGAGCTTCTGTGTAATAGGCA 3162  
 Db 421 AACTCCCAAGAGCAAAATCCACATTCCTCTTGAGTTCTCGAGCTTCTGTGTAATAGGCA 480  
 QY 3163 GCTGCTGCTATGCGGTAGATCAATCATGATCTGAGGACCATTCATGGAAGCTGCTAAATA 3222  
 Db 481 GCTGCTGCTATGCGGTAGATCAATCATGATCTGAGGACCATTCATGGAAGCTGCTAAATA 540  
 QY 3223 GCCTAGTCTGGGAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTTAACTA 3282  
 Db 541 GCCTAGTCTGGGAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTTAACTA 600  
 QY 3283 GCTTTGGTTCTTTCAGCCCTCTAAAGCATAGGCTTAGCCCTGAGGCTTCTCTGGGCTT 3342  
 Db 601 GCTTTGGTTCTTTCAGCCCTCTAAAGCATAGGCTTAGCCCTGAGGCTTCTCTGGGCTT 660  
 QY 3343 TCTCTGTGTGTAGTTTGTAAACACATATAGCATCTGTAAAGTCCAGTGTCCATGAA 3402  
 Db 661 TCTCTGTGTGTAGTTTGTAAACACATATAGCATCTGTAAAGTCCAGTGTCCATGAA 720  
 QY 3403 ACCTTCCCAATGCGGTAGTCTGCACTATATATAGTTTTCGAAAGCAGGTTCTCTGTC 3462  
 Db 721 ACATTTCCCAATGCGGTAGTCTGCACTATATATAGTTTTCGAAAGCAGGTTCTCTGTC 780  
 QY 3463 CTGCTAAACAAGCCCACTGAGGACCTCTGAATGTCTTCCCTTACACCTATGTTTTTAAG 3522  
 Db 781 CTGCTAAACAAGCCCACTGAGGACCTCTGAATGTCTTCCCTTACACCTATGTTTTTAAG 840  
 QY 3523 TAGTCAAACTTCAAGAAACAATCTAAACAAGTTTCTGTGCAATATGTTGTGGAAGCTT 3582  
 Db 841 TAGTCAAACTTCAAGAAACAATCTAAACAAGTTTCTGTGCAATATGTTGTGGAAGCTT 900  
 QY 3583 TATTGTATTTAGTAGGCTCTATATATGCAATTTAACTTTGTTTCTAACTCTGATTTCTT 3642  
 Db 901 TATTGTATTTAGTAGGCTCTATATATGCAATTTAACTTTGTTTCTAACTCTGATTTCTT 960  
 QY 3643 CCTTTTCGATATTTGATGTAATTAAGAAATTAAGTGAAGAAAAA 3692  
 Db 961 CCTTTTCGATATTTGATGTAATTAAGAAATTAAGTGAAGAAAAA 1010

RESULT 9  
 AAX39840  
 ID AAX39840 standard; DNA; 1017 BP.  
 XX  
 AC AAX39840;  
 XX  
 DT 02-JUL-1999 (first entry)  
 XX  
 DE Gastric cancer associated gene.  
 XX  
 KW Cancer associated antigen; diagnosis; research; treatment; human;  
 KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;  
 KW prostate cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO9904265-A2.  
 XX  
 PD 28-JAN-1999.  
 XX  
 PF 15-JUL-1998; 98WO-US014679.  
 XX  
 PR 17-JUL-1997; 97US-00896164.  
 PR 10-OCT-1997; 97US-0061599P.  
 PR 10-OCT-1997; 97US-0061765P.  
 PR 10-OCT-1997; 97US-00948705.  
 PR 11-OCT-1997; 97GB-00021697.  
 PR 22-JUN-1998; 98US-00102322.  
 XX  
 PA (LUDW-) LUDWIG INST CANCER RES.

Db 618 TCCCTAGATACAGAGTCTTATACCCACACACCTGTCTACAGTGTCAATTAAGTG 677  
Qy 806 -ACGTCCTGTGAAAAAACAATAAACC 837  
Db 678 CACGTCCTGTGAAAAAACAATAAACC 710

RESULT 15  
CD175286  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CD175286 712 bp mRNA linear EST 19-MAY-2003  
AGENCOURT\_13978429 NIH\_MGC\_172 Homo sapiens cDNA 5', mRNA sequence.  
CD175286  
CD175286.1 GI:30859713  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue Procurement: Dr. Jamie Thompson, University of WI  
cDNA Library Preparation: Gina Zastrow-Hayes  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILLNL at:  
<http://image.llnl.gov>  
Plate: NDKM49 row: m column: 06  
High quality sequence start: 31  
High quality sequence stop: 712.

FEATURES  
Location/Qualifiers  
1..712  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, WA01, passage 39"  
/lab\_host="DH10B TcrA"  
/clone\_lib="NIH MGC 172"  
/note="Vector: pDONR201; Site 1: attP2; Site 2: attP1;  
LIBR\_PRIMING - Oligo dT; METHOD - full-length enriched;  
Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"

ORIGIN  
Query Match 60.1%; Score 666.4; DB 14; Length 712;  
Best Local Similarity 98.3%; Pred. No. 9.6e-125;  
Matches 684; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

Qy 143 GAGAACTGAACGACAGGGGAAAGAGTCTCAGTACGACCGTCCAGCATCCGGACA 202  
Db 17 GAGTAGAAAAACGACAGGGGAAAGAGTCTCAGTACGACCGTCCAGCATCCGGACA 76  
Qy 203 CCACAGCGGCCCTTCGCTCCAGCGAAAAACACACTTCTCAAACTTCACCTCAACACTT 262  
Db 77 CCACAGCGGCCCTTCGCTCCAGCGAAAAACACACTTCTCAAACTTCACCTCAACACTT 136  
Qy 263 CTTTCCCAAGCCAGAAATGCAAGAGGAAACATGAGTGGTGTCTGGGGGACC 322  
Db 137 CTTTCCCAAGCCAGAAATGCAAGAGGAAACATGAGTGGTGTCTGGGGGACC 196  
Qy 323 CCCCAGCAGCTCTTCCAGGTCCACCGTGTACATCCACGAGACCTCCGTGCC 382  
Db 197 CCCCAGCAGCTCTTCCAGGTCCACCGTGTACATCCACGAGACCTCCGTGCC 256  
Qy 383 CGACCATGTCTGTCTGCTTCAACACCTCTTCTTTGAATGGTGTCTGTGGGCTT 442  
Db 257 CGACCATGTCTGTCTGCTTCAACACCTCTTCTTTGAATGGTGTCTGTGGGCTT 316  
Qy 443 CATAGCAATCGCCTACTCCGTGAACTCTAGGACAGGAAGATGGTTGGCGACGTGACCGG 502

Db 317 CATAGCAATCGCCTACTCCGTGAAGTCTAGGGACAGGAAGATGGTTGGCGACGTGACCGG 376  
Qy 503 GGCCCAAGCCCTATGCTCCACCGCAAGTCCCTGAACATCTGGGCCCTGATTTCTGGCAT 562  
Db 377 GGCCCAAGCCCTATGCTCCACCGCAAGTCCCTGAACATCTGGGCCCTGATTTCTGGCAT 436  
Qy 563 CCTCATGACCAATGGATTCAATCCTGTTACTTGGTATTTCGGCTCTGTGACAGTCTTACCATAT 622  
Db 437 CCTCATGACCAATGGATTCAATCCTGTTACTTGGTATTTCGGCTCTGTGACAGTCTTACCATAT 496  
Qy 623 TATGTTACAGATAATACAGGAAAAACGGGTTACTAGTAGCCGCCCATAGCCTGCAACCT 682  
Db 497 TATGTTACAGATAATACAGGAAAAACGGGTTACTAGTAGCCGCCCATAGCCTGCAACCT 556  
Qy 683 TTGCACCTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGTGCCCTTGGCCCTTGG 742  
Db 557 TTGCACCTCCACTGTGCAATGCTGGCCCTGCACACTGGGGCTGTGTGCCCTTGGCCCTTGG 616  
Qy 743 TCCTGCCCCCTAGATACAGCAGTATTATACCCACACACACTGTCTACACTGACATTCAATAAA 802  
Db 617 TCCTGCCCCCTAGATACAGCAGTATTATACCCACACACACTGTCTACAGTGTCAITCAATAAA 676  
Qy 803 GTG-ACGTGCTTGTGAAAAAACAATAAACC 837  
Db 677 GTGACGCTGCTTGTGAAAAAACAATAAACC 712

Search completed: March 11, 2004, 14:23:57  
Job time : 1840.29 secs

XX Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;  
 PI O'hare M, Obata Y, Pfreundschuh M, Turesi O, Sahin U;  
 XX WPI; 1999-132448/11.  
 XX New isolated cancer associated nucleic acids and polypeptides - isolated  
 PT using sera from cancer patients, used to develop products for the  
 PT diagnosis, monitoring or treatment of cancers.  
 XX Claim 67; Page 567; 787pp; English.  
 XX The invention relates to a method for diagnosing a disorder characterised  
 CC by expression of a human cancer associated antigen precursor coded for by  
 CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a  
 CC biological sample isolated from a subject with an agent that specifically  
 CC binds to the NAM, an expression product or a fragment of an expression  
 CC product complexed with an HLA molecule; and (b) determining the  
 CC interaction between the agent and the NAM or the expression product as a  
 CC determination of the disorder. The products and methods can be used in  
 CC the diagnosis, monitoring, research, or treatment of conditions  
 CC characterised by the expression of various cancer associated antigens.  
 CC The invention provides nucleic acid sequences and encoded polypeptides  
 CC which are cancer associated antigen precursors expressed in human breast  
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and  
 CC lung cancer  
 XX Sequence 1017 BP; 268 A; 257 C; 274 G; 211 T; 0 U; 7 Other;

QY 612 CTGCCATCAATATTACACGTATCCAGACCCAGAGAACCAACAGTTGCTTAGGGACTCT 671  
 DB 613 CTGCCATCAATATTACACGTATCCAGACCCAGAGAACCAACAGTTGCTTAGGGACTCT 672  
 QY 672 GTGTTGATGCTCTCATTTGAACCTGCTGATGAATAATGCTGTAATGGAACCTCAGCTTCCAG 731  
 DB 673 GTGTTGATGCTCTCATTTGAACCTGCTGATGAATAATGCTGTAATGGAACCTCAGCTTCCAG 732  
 QY 732 AGTTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCTCTGAGAAAGTGTGCCCTGGAGG 791  
 DB 733 AGTTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCTCTGAGAAAGTGTGCCCTGGAGG 792  
 QY 792 ATGAACCTGATCCAGATGAGCTGAGACCGAGTGGACTGTAA-CGGTGTCTCTGTGCC 850  
 DB 793 ATGAACCTGATCCAGATGAGCTGAGACCGAGTGGACTGTAA-CGGTGTCTCTGTGCC 852  
 QY 851 TGTGGAATTTGGTCTGTGACAGCCATGACCTGTGACGGAAGAAATCAGAAAGGGGCCCCAG 910  
 DB 853 TGTGGAATTTGGTCTGTGACAGCCATGACCTGTGACGGAAGAAATCAGAAAGGGGCCCCAG 911  
 QY 911 ACCCAGACAGAGAGAGATGACAGATATGTCAGGAGCTCCAAAGCATCAGGAACA 970  
 DB 912 ACCCAGACAGAGAGAGATGACAGATATGTCAGGAGCTCCAAAGCATCAGGAACA 970  
 QY 971 GCTGAAAAGACCAAGAGAGTGGACCAACCAAGAGATCTAATGAGGAGGC 1018  
 DB 971 CTGAAAAGACCAAGAGAGTGGACCAACCAAGAGAGTGGACCAACCAAGAGAGGC 1017

RESULT 10  
 AAT68831  
 ID AAT68831 standard; cDNA to mRNA; 926 BP.  
 XX AC AAT68831;  
 XX DT 04-FEB-1998 (first entry)  
 XX DE cDNA for follistatin related protein.  
 XX KW Rheumatoid arthritis; auto-antigen; FRP; diagnosis; prediction;  
 XX KW synovial cell; follistatin related protein; ss.  
 XX OS Homo sapiens.  
 FH Key Location/Qualifiers  
 CDS 1..926  
 FT /\*tag= a  
 FT /product= "follistatin related protein"  
 FT /transl\_except= (pos: 523..524, aa:Asn)  
 FT /note= "this codon has apparent 1 nucleotide deletion,  
 FT which alters reading frame"  
 XX WO9717441-A1.  
 XX 15-MAY-1997.  
 XX 06-NOV-1996; 56WO-JP003250.  
 XX 07-NOV-1995; 95JP-00286957.  
 XX (KANF ) KANEKA CORP.  
 XX Osaki S, Tanaka M, Kishimura M, Nakao K, Osakada F;  
 XX WPI; 1997-381030/25.  
 XX P-PSDB; AAW17860.  
 XX Auto-antigen from synovial cells of rheumatoid arthritis patients - binds  
 PT to antibodies present in these patients, for diagnosis and prediction of  
 PT the disease.  
 XX Claim 9; Page 40-41; 61pp; Japanese.  
 XX







1959	GAAGAGCATCTTCATATAAAAACTTTTCAGATGCCAAACATTAATTTGTTTATTATTATTC	2019
248	GAAGAGCATCTTCATATAAAAACTTTTCAGATGCCAAACATTAATTTGTTTATTATTATTC	307
2019	TGAGAACTTGAGGCAATCAGTATTCCTCCCAAGGATGGCGACAAGGGCAGCCAAAGCAGGGCT	2078
308	TGAGAACTTGAGGCAATCAGTATTCCTCCCAAGGATGGCGACAAGGGCAGCCAAAGCAGGGCT	367
2079	TAGGATATCCCAAGCCTACCAATATGCTCAATCGACCTAATAGGAGGGTGAGTTCGGCCCTG	2138
368	TAGGATATCCCAAGCCTACCAATATGCTCAATCGACCTAATAGGAGGGTGAGTTCGGCCCTG	427
2139	TCTCTTCTTTTTCGGAACCTCAGTTTCTCTCAGTGAGCTGGTAAGAAATGCACCTAACCTTTT	2198
428	TCTCTTCTTTTTCGGAACCTCAGTTTCTCTCAGTGAGCTGGTAAGAAATGCACCTAACCTTTT	487
2199	TGATTTGATAAGTTATATAAATTCGTGGTTCGTGATCATTTGGTCAGAGGGGAGATAGGTTC	2258
488	TGATTTGATAAGTTATATAAATTCGTGGTTCGTGATCATTTGGTCAGAGGGGAGATAGGTTC	547
2259	CTGTGATTTTTTCCCTTCTTCTCTATAGAAATAAATGAAATCTTGTTACTAGAACAGAAATG	2318
548	CTGTGATTTTTTCCCTTCTTCTCTATAGAAATAAATGAAATCTTGTTACTAGAACAGAAATG	607
2319	TCAGATGCCCAAAAAACAAGATACCAAGATTTTGATCTTCAGCCTGTAGACCTACAGGTCGT	2378
608	TCAGATGCCCAAAAAACAAGATACCAAGATTTTGATCTTCAGCCTGTAGACCTACAGGTCGT	667
2379	GCTATGATATGGAGTCCTCATGGTAAAGCAGGAGAGAGTGGGAAAGAGAACCAACCCCA	2438
668	GCTATGATATGGAGTCCTCATGGTAAAGCAGGAGAGAGTGGGAAAGAGAACCAACCCCA	727
2439	CTCTGCTCTTCATATTTGCATTTTCAATGTTTAACTCCGGCTGGAAATAGAAAGCATTCCTT	2498
728	CTCTGCTCTTCATATTTGCATTTTCAATGTTTAACTCCGGCTGGAAATAGAAAGCATTCCTT	787
2499	TAGAGATGAGGATATAAAGAAAGTTTCAGATTCACACAGGGGAA-CAAAATGGAGATTAA	2557
788	TAGAGATGAGGATATAAAGAAAGTTTCAGATTCACACAGGGGAAAGAAATGGAGATTAA	847
2558	TCCTAAAACTGTGACTTTGGGAGGTCAGTCAT	2589
848	TCCTAAAACTGTGACTTTGGGAGGTCAGTCAT	879

RESULT 13

RESULT 13  
AAY398A2

AX39842  
IR AX39842 standard: DNA: 881 BP.

ID: AAX39842

1000

4C AAX39842;



DT 02-JUL-19

25

Gastric c

XX  
ET  
TITAN

DECEMBER 1964

Cancer as

breast ca

KW prostate



SC Homo sapi

4

PN WO9904265

XX  
XX  
CONFIDENTIAL

28-JAN-19

XX  
ED  
28-JAN-82

XX  
15-III-19

15-JUL-19

**X**

PR 17-JUL-19

PR 10-OCT-19

PR 10-OCT-19

PR 10-OCT-19

PR 11-OCT-19

PR 11-OCT-19

PR 22-JUN-19

XX	(LUDW-) LUDWIG INST CANCER RES.
XX	
PI	Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;
PI	O'hare M, Obata Y, Pfreundschuh M, Tureci O, Sahin U;
XX	
DR	WPI; 1999-132448/11.
XX	
PT	New isolated cancer associated nucleic acids and polypeptides - isolated
PT	using sera from cancer patients, used to develop products for the
PT	diagnosis, monitoring or treatment of cancers.
XX	
PS	Claim 67; Page 568; 787pp; English.
XX	
CC	The invention relates to a method for diagnosing a disorder characterised
CC	by expression of a human cancer associated antigen precursor coded for by
CC	a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC	biological sample isolated from a subject with an agent that specifically
CC	binds to the NAM, an expression product or a fragment of an expression
CC	product complexed with an HLA molecule, and (b) determining the
CC	interaction between the agent and the NAM or the expression product as a
CC	determination of the disorder. The products and methods can be used in
CC	the diagnosis, monitoring, research, or treatment of conditions
CC	characterised by the expression of various cancer associated antigens.
CC	The invention provides nucleic acid sequences and encoded polypeptides
CC	which are cancer associated antigen precursors expressed in human breast
CC	cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC	lung cancer
XX	
SQ	Sequence 881 BP; 225 A; 219 C; 231 G; 197 T; 0 U; 9 Other;
	Query Match            20.5%; Score 757.2; DB 2; Length 881;
	Best Local Similarity 95.8%; Pred. No. 7.7e-184;
	Matches 815; Conservative 0; Mismatches 32; Indels 4; Gaps 4;
Qy	12 CTCCGACCTCGGTATACAGTCTCGTCGGCGGCTCTGCCCGCGCCCCAGGAGACTGG 71
Db	
	13 CTCGCACCTCGGTATACAGTCTCGTCGGCGGCTCTGCCCGCGCCCCAGGAGACTGG 72
Qy	72 ACCAGACCAGATGTGGAACCGTGCTCGCGCTCGCGCTCGCGCTGGTGCGGTGGCTCGCCT 131
Db	
	73 ACCAGACCAGATGTGGAACCGTGCTCGCGCTCGCGCTCGCGCTGGTGCGGTGGCTCGCCT 132
Qy	132 GGGTCGGCGCGGAGAGAGACTAAGGCAAAATCCRAGATCTGTGCCAAATGTTGTTGTG 191
Db	
	133 GGGTCGGCGCGGAGAGAGACTAAGGCAAAATCCRAGATCTGTGCCAAATGTTGTTGTG 192
Qy	192 GAGCGGCGCGGGAATGTGCAGTGCAGTGCAGAGAAAAGGGGAACCCACTGCTCTTCGATTGAGC 251
Db	
	193 GAGCGGCGCGGGAATGTGCAGTGCAGTGCAGAGAAAAGGGGAACCCACTGCTCTTCGATTGAGC 252
Qy	252 ARTGCAAACTCACAGAGCGCTGTGTGTGCAGTATGTGCRAGACCTACTCTCAACCACT 311
Db	
	253 ARTGCAAACTCACAGAGCGCTGTGTGTGCAGTATGTGCRAGACCTACTCTCAACCACT 312
Qy	312 GTGAACCTGCATCGAGATGCCTGCGCTCACTGGATGCAAAATCCAGGTTGATTACGATGGAC 371
Db	
	313 GTGAACCTGCATCGAGATGCCTGCGCTCACTGGATGCAAAATCCAGGTTGATTACGATGGAC 372
Qy	372 ACTGCAAGAAGAGAAATCGTAAGTTCATCTGCCAGCCAGTTGTTGTCTATCAGTCCA 431
Db	
	373 ACTGCAAGAAGAGAAATCGTAAGTTCATCTGCCAGCCAGTTGTTGTCTATCAGTCCA 432
Qy	432 ACCGTGATGAGCTCCGACGTCGCATCATCCAGTGGCTGGAGCTGAGATCATCTCAGATG 491
Db	
	433 ACCGTGATGAGCTCCGACGTCGCATCATCCAGTGGCTGGAGCTGAGATCATCTCAGATG 492
Qy	492 GCTGGTTCTTAAGGCGAGCACTACAGTGAATCCTAGCCAGTAGTATTTTAAGAACCTTG 551
Db	
	493 GCTGGTTCTTAAGGCGAGCACTACAGTGAATCCTAGCCAGTAGTATTTTAAGAACCTTG 552
Qy	552 ATAATGGTGATTCTTCGCTCGGACTCCAGTGAATTCCTGAAGTTTGTGGAAACAGAAATGAAA 611

553 ATAATGGTGATTCCTCGCTGGAGCTCCAGTGAAATTCCTGGAAGTTTGTGGAAACAGAAATGAAA 611

612 CTGCATCAATATTACAACTGATCCAGACGAGGAGAACAAAGTTTGCCTTAGGGGACTCT 671

613 CTGCATCAATATTACAACTGATCCAGACGAGGAGAACAAAGTTTGCCTTAGGGGACTCT 672

672 GTGTTGATGCTCTCAATTGAACTGCTGATGAAATGCTGATGAAACTCAGCTTCCAG 731

673 GGTTTGATGCTCTCAATTGAACTGCTGATGAAATGCTGATGAAACTTAGCTTTCAAG 732

732 -AGTTTCTCAAGTGCC-TCAACCCCATCTTTCAACCCCTCTGAGAGAAAGTGTGCCCTGGA 789

733 AAGTTTCTCAAGNGCCTTNAACCCATCTTTTAAACCTTTCTTGAGAGAAATGTGCCCTTGA 792

790 GGATGAACAGTATG-CAGATGGAGC-TGAGACCGAGTGGACTGTAACCGCTGTGTCTGT 847

793 GGATGAACAGTATGCCANATGGAGCTTGAANCGAGTGGACTGTAAACCGTTGGNCTGG 852

848 GCCTGTGGAAA 858

853 GNCCTGGGGAA 863

RESULT 14

ABZ34644

ID ABZ34644 standard; cDNA; 930 BP.

XX AC ABZ34644;

XX AC

XX AC

DT DT 04-FEB-2003 (first entry)

XX XX

DE DE Coding sequence SEQ ID 2, upregulated in osteogenesis.

XX XX Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;

KW KW osteoporosis; bone disease; downregulator; human; ss.

XX XX

XX XX Homo sapiens.

OS OS

XX XX WO200281745-A2.

PN PN

XX XX

PD PD 17-OCT-2002.

XX XX

PF PF 05-APR-2002; 2002WO-IB002211.

XX XX

PR PR 05-APR-2001; 2001US-0281400P.

XX XX

PA PA (AVET ) AVENTIS PHARMA SA.

XX XX

PI PI Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;

PI PI Connolly T, Jackson A, Bushnell SE, Rawadi G;

XX XX

DR DR WPI; 2003-059567/05.

XX XX

PT PT Novel isolated nucleic acid upregulated/downregulated in osteogenesis,

PT PT useful for bone disease therapy in subject.

XX XX

PS PS Claim 26; Page 84; 237pp; English.

XX XX

CC CC The present invention relates to novel nucleotide sequences, which are

CC CC differentially expressed in models of osteogenesis upon being put in

CC CC contact with a stimulator of osteogenesis. The present sequence is one

CC CC such sequence. This sequence can be used for diagnosing osteoporosis/bone

CC CC disease in a patient, promoting osteogenesis and/or preventing

CC CC osteoporosis/bone disease. The present sequence encodes a secreted

XX XX protein

XX XX

SQ SQ Sequence 930 BP; 263 A; 236 C; 247 G; 184 T; 0 U; 0 Other;

Query Match 19.8%; Score 729.4; DB 7; Length 930;

Best Local Similarity 87.8%; Pred. No. 1.1e-176;

Matches 796; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

Oy 103 GCTCGCCTCGCGCTGGTGGCGTCCCTCGCTGGTGGCGCGCGAGGAGCTAAGGACAA 162

X Cancer associated antigen; diagnosis; research; treatment; human;  
W breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;  
M prostate cancer; ss.  
X Homo sapiens.  
S WO9904265-A2.  
N 28-JAN-1999.  
X 15-JUL-1998; 98WO-US014679.  
X 17-JUL-1997; 97US-00896164.  
R 10-OCT-1997; 97US-0061599P.  
R 10-OCT-1997; 97US-0061765P.  
R 10-OCT-1997; 97US-00948705.  
R 11-OCT-1997; 97GB-00021697.  
R 22-JUN-1998; 98US-00102322.  
X (LUDW-) LUDWIG INST CANCER RES.  
X Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;  
I O'hare M, Obata Y, Pfreundschuh M, Tureci O, Sahin U;  
X WPI; 1999-132448/11.  
X New isolated cancer associated nucleic acids and polypeptides - isolated  
X using sera from cancer patients, used to develop products for the  
X diagnosis, monitoring or treatment of cancers.  
X Claim 67; Page 568; 787pp; English.  
X The invention relates to a method for diagnosing a disorder characterised  
X by expression of a human cancer associated antigen precursor coded for by  
X a nucleic acid molecule (NAM). The method comprises: (a) contacting a  
X biological sample isolated from a subject with an agent that specifically  
X binds to the NAM, an expression product or a fragment of an expression  
X product complexed with an HLA molecule; and (b) determining the  
X interaction between the agent and the NAM or the expression product as a  
X determination of the disorder. The products and methods can be used in  
X the diagnosis, monitoring, research, or treatment of conditions  
X characterised by the expression of various cancer associated antigens.  
X The invention provides nucleic acid sequences and encoded polypeptides  
X which are cancer associated antigen precursors expressed in human breast  
X cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and  
X lung cancer  
X Sequence 810 BP; 234 A; 186 C; 206 G; 177 T; 0 U; 7 Other;  
Query Match 19.7%; Score 729.2; DB 2; Length 810;  
Best Local Similarity 96.4%; Pred. No. 1.2e-176;  
Matches 784; Conservative 0; Mismatches 25; Indels 4; Gaps 4;  
Y 207 GTCCAGTTCACAGAGAAGGGGAACCCACCTGTCTTGCATTGAGCAATCCAAACCTCACA 266  
b 1 GTCCAGTTCACAGAGAAGGGGAACCCACCTGTCTTGCATTGAGCAATCCAAACCTCACA 60  
Y 267 AGAGGCTGTGTGTCGAGTAATGGCAAGACCTACCTCAACCACTGTGAACCTCGAG 326  
b 61 AGAGGCTGTGTGTCGAGTAATGGCAAGACCTACCTCAACCACTGTGAACCTCGAG 120  
Y 327 ATCCCTGCCTCACTGGATCCAAATCCAGTTGATTACGATGACACCTGCAAGAGAAGA 386  
b 121 ATCCCTGCCTCACTGGATCCAAATCCAGTTGATTACGATGACACCTGCAAGAGAAGA 180  
Y 387 AATCCGTGAGTCCATCTGCCAGCCAGTTGTTTGTCTATCAGTCCACCGTGATGAGTCC 446  
b 181 AATCCGTGAGTCCATCTGCCAGCCAGTTGTTTGTCTATCAGTCCACCGTGATGAGTCC 240  
Y 447 GACGTGCGATCATCCAGTGGCTGGAAAGCTGAGATCATCTCAGATGCGTGTCTCTAAAG 506  
b 241 GACGTGCGATCATCCAGTGGCTGGAAAGCTGAGATCATCTCAGATGCGTGTCTCTAAAG 300

QY 507 GCAGCACTACAGTGAATCCTAGACAGAGTATTTTAAAGACTTTTGATATGGTGATCTC 566  
Db 301 GCAGCACTACAGTGAATCCTAGACAGAGTATTTTAAAGACTTTTGATATGGTGATCTC 360  
QY 567 GCCTGGACTCCAGTGAATCCTTGAAGTTTGTGGACAGAGTAAACTGCCATCAATATTA 626  
Db 361 GCCTGGACTCCAGTGAATCCTTGAAGTTTGTGGACAGAGTAAACTGCCATCAATATTA 420  
QY 627 CAACGTATCCAGACCCAGGAGAACCAAGTTGCTTAGGGGACTCTGTGTGATGCTCTCA 686  
Db 421 CAACGTATCCAGACCCAGGAGAACCAAGTTGCTTAGGGGACTCTGTGTGATGCTCTCA 480  
QY 687 TTGAACCTGTCTGATGAAATGCTGATTTGAAATCTCAGCTTCCAGAGTCTTCTCAAGTCC 746  
Db 481 TTGAACCTGTCTGATGAAATGCTGATTTGAAATCTCAGCTTCCAGAGTCTTCTCAAGTCC 540  
QY 747 TCAACCCATCTTTCAACCCCTCTGAGAGAGAGTGTGCTTGGAGGATGAAACGTATGCGAG 806  
Db 541 TCAACCCATCTTTCAACCCCTCTGAGAGAGAGTGTGCTTGGAGGATGAAACGTATGCGAG 600  
QY 807 ATGGAGCTGAGACCCAGGTGGACTGTAA - CCGCTGTGTCTGTGCTGTGAAATTTGGGTC 865  
Db 601 ATGGAGCTGAGACCCAGGTGGACTGTAA - CCGCTGTGTCTGTGCTGTGAAATTTGGGTC 660  
QY 866 TGTACAGCCATGACCTGTGTGACCGAAGAGATTCAGAGGGGGCCCGAGACCCAGACAGAGGAG 925  
Db 661 TGT - CAGCCATGACCTGTGTGACNGAAGAGATTCAGAGGGGGCCCGAGACCCAGACNGAGGAN 719  
QY 926 GAGATGACCCAGATATGTCCAGGAGCTCCAAAGCATCAGGAAACAGCTGAAAGACCAAG 985  
Db 720 GAGATGACNGATATGTCCAGGAGCTCC - AAAGCTTAGGAAACAGCTTGAAGAACCCAG 778  
QY 986 AGAGTGAAGCCAAAGAGATCTTAATGAGGAGGC 1018  
Db 779 AGAGGAGGCCCCCAA - AGAGATTAATGAGGAGGC 810

Search completed: March 10, 2004, 10:59:23  
Job time : 904.804 secs



Result No.	Score	Query		ID	Description
		Match	Length		
1	3693	100.0	3693	4	US-09-976-594-908
2	892.8	24.2	896	4	Sequence 908, App
3	619.4	16.8	621	4	Sequence 15155, A
4	619.4	16.8	621	4	Sequence 209, App
5	619.4	16.8	621	4	Sequence 209, App
6	619.4	16.8	621	4	Sequence 209, App
7	477.2	12.9	708	3	Sequence 34, App
8	336	9.1	336	4	Sequence 34, App
9	336	9.1	336	4	Sequence 315, App
10	336	9.1	336	4	Sequence 315, App
11	336	9.1	336	4	Sequence 315, App
12	70.2	1.9	1355	4	Sequence 315, App
13	70.2	1.9	1392	4	Sequence 7, Appli
14	66.8	1.8	2949	4	Sequence 1, Appli
15	53.2	1.4	7218	1	Sequence 3, Appli
16	50	1.4	7218	1	Sequence 14, Appl
17	47	1.3	1771	4	Sequence 14, Appl
18	47	1.3	1771	4	Sequence 158, App
19	47	1.3	1771	4	Sequence 36, Appl
20	47	1.3	1771	4	Sequence 158, App
21	45.8	1.2	11050	4	Sequence 158, App
22	45.6	1.2	5852	1	Sequence 86, Appl
23	45	1.2	6843	2	Sequence 2, Appli
24	44.2	1.2	6801	4	Sequence 1, Appli
25	44.2	1.2	11049	4	Sequence 62, Appl
26	43.8	1.2	1813	3	Sequence 22, Appl
27	43.8	1.2	5152	4	Sequence 3, Appli





2521 TTTTCAGATTCAACAGGGGGAAGAAATGGAGATTAAATCCTAAATCTGTGACTTGGGGAG 2580  
2521 TTTTCAGATTCAACAGGGGGAAGAAATGGAGATTAAATCCTAAATCTGTGACTTGGGGAG 2580  
2581 GTGAGTCAATTTACAGTTAGTCTGTGTCCTTTCAGACTCTGTGATTAATTAACCCACTCAC 2640  
2581 GTGAGTCAATTTACAGTTAGTCTGTGTCCTTTCAGACTCTGTGATTAATTAACCCACTCAC 2640  
2641 TACCTCTGTTTCAGATGCAATTTGGAATACCAAGAAATAAATCCTTTGACATAAGATCTCAAT 2700  
2641 TACCTCTGTTTCAGATGCAATTTGGAATACCAAGAAATAAATCCTTTGACATAAGATCTCAAT 2700  
2701 TGCAGAAAGCAGATTAAGAACCATCAGAAAGAAATTAATTTAGTGTGTAATGCAAGGCAA 2760  
2701 TGCAGAAAGCAGATTAAGAACCATCAGAAAGAAATTAATTTAGTGTGTAATGCAAGGCAA 2760  
2761 CTGTGAGAAATCTGTGTGGCCAAAATAGAAATCCTCTAGTCTTTCTGTTCTCAATTTGA 2820  
2761 CTGTGAGAAATCTGTGTGGCCAAAATAGAAATCCTCTAGTCTTTCTGTTCTCAATTTGA 2820  
2821 AAGGAGAAATTCACATTTGTTTGTAGCAATTTCAAGCTTTTATGTATCCATCCCATCTAAA 2880  
2821 AAGGAGAAATTCACATTTGTTTGTAGCAATTTCAAGCTTTTATGTATCCATCCCATCTAAA 2880  
2881 ACTCTTCAAACTCCACTTTGTTTGTAGCAATTTCAAGCTTTTATGTATCCATCCCATCTAAA 2940  
2881 ACTCTTCAAACTCCACTTTGTTTGTAGCAATTTCAAGCTTTTATGTATCCATCCCATCTAAA 2940  
2941 ACTCACAGCAGCAGCCCTTAATCAAGAGTTTACCAGCCCTTGACACTATGGAGGAGG 3000  
2941 ACTCACAGCAGCAGCCCTTAATCAAGAGTTTACCAGCCCTTGACACTATGGAGGAGG 3000  
3001 GCAGAGTACACCAATTTGTTTAAAGCAAGAAACCAAGAGTCTCTTCACTAGTCAATTTA 3060  
3001 GCAAGAGTACACCAATTTGTTTAAAGCAAGAAACCAAGAGTCTCTTCACTAGTCAATTTA 3060  
3061 GAACATGGTTATCATCAAGACTACTCTACCTGCAACATTTGAATCCTCCCAAGAGCAATC 3120  
3061 GAACATGGTTATCATCAAGACTACTCTACCTGCAACATTTGAATCCTCCCAAGAGCAATC 3120  
3121 CACATTCCTCTGTAGTCTGCGAGCTTCTGTGTAATAGGCGAGCTGCGCTATGCGGTA 3180  
3121 CACATTCCTCTGTAGTCTGCGAGCTTCTGTGTAATAGGCGAGCTGCGCTATGCGGTA 3180  
3181 GAATCATGATCTGAGGACCAATTCATGGAAGCTGTCTAAATAGCTAGTCTGGGGAGTCT 3240  
3181 GAATCATGATCTGAGGACCAATTCATGGAAGCTGTCTAAATAGCTAGTCTGGGGAGTCT 3240  
3241 TCCATAAAGTTTTCGATGGAGCAACAAACAGGATTAACCTAGGTTTGGTTCCTTCAGCC 3300  
3241 TCCATAAAGTTTTCGATGGAGCAACAAACAGGATTAACCTAGGTTTGGTTCCTTCAGCC 3300  
3301 CTCTAAAGAGCATAGGGCTTAGCTTGCAGGCTTCTCTGGGCTTTCTCTGTGTGTAGTTT 3360  
3301 CTCTAAAGAGCATAGGGCTTAGCTTGCAGGCTTCTCTGGGCTTTCTCTGTGTGTAGTTT 3360  
3361 TGTAAACACTATAGCATCTGTATAGATCCAGTGTCCATGGAACCTTCCACATGCGCGT 3420  
3361 TGTAAACACTATAGCATCTGTATAGATCCAGTGTCCATGGAACCTTCCACATGCGCGT 3420  
3421 ACTCTGGAATATACAGTTTGTGAAAGCAGGGTCTCTGCGCTGTCTGCTGTGTGTAGTTT 3480  
3421 ACTCTGGAATATACAGTTTGTGAAAGCAGGGTCTCTGCGCTGTCTGCTGTGTGTAGTTT 3480  
3481 GGACCAAGCTGAATGTCTTCTTTACACCTATGTTTAAAGTAGTCAAACTTCAAGAAA 3540  
3481 GGACCAAGCTGAATGTCTTCTTTACACCTATGTTTAAAGTAGTCAAACTTCAAGAAA 3540  
3541 CAATCTAAACAGTTTCTGTTGCAATGTGTTGTAAGTCTGTTGTAATTTAGTAGG 3600  
3541 CAATCTAAACAGTTTCTGTTGCAATGTGTTGTAAGTCTGTTGTAATTTAGTAGG 3600  
3601 TTCTATATTGCAATTAACCTGTTTGTGTAATCTCTGATTTCTTCTTTTCGATACTATTG 3660

Db 3601 TTCTATATTGCAATTAACCTGTTTGTGTAATCTCTGATTTCTCTTTTCGATACTATTG 3660  
QY 3661 ATGAATAAAGAAATTAAGAGTGAATAAAGTGAATAAAGTGAATAAAGTGAATAAAGTGA 3693  
Db 3661 ATGAATAAAGAAATTAAGAGTGAATAAAGTGAATAAAGTGAATAAAGTGAATAAAGTGA 3693  
RESULT 2  
US-09-621-976-15155  
; Sequence 15155, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 15155  
; LENGTH: 896  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-15155  
Query Match 24.2%; Score 892.8; DB 4; Length 896;  
Best Local Similarity 99.8%; Pred. No. 2.1e-243;  
Matches 894; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2668 CCAAGATTAAATCCTTGACATAAGATCTCATTTGCAAGAGAGAGATTAAGAGCATTAG 2727  
Db 1 CCAAGATTAAAGCCTTGACATAAGATCTCATTTGCAAGAGAGAGATTAAGAGCATTAG 60  
QY 2728 AAGGAAATTAATTTAGGTTGTAATGCAAGGCAACTGTGAGAAACTGTGTGCCAAAATA 2787  
Db 61 AAGGAAATTAATTTAGGTTGTAATGCAAGGCAACTGTGTGCCAAAATA 120  
QY 2788 GAATTCCTCTAGTCTTCTGTTCTCATTTGAAAGAGAGAAATTCCTCACTTTGTTAGCA 2847  
Db 121 GAATTCCTCTAGTCTTCTGTTCTCATTTGAAAGAGAGAAATTCCTCACTTTGTTAGCA 180  
QY 2848 TTCAAAGCTTTTATGATCCATCCATCTAAATCTTCAAACTCCACTTTGTTCAAGTCT 2907  
Db 181 TTCAAAGCTTTTATGATCCATCCATCTAAATCTTCAAACTCCACTTTGTTCAAGTCT 240  
QY 2908 GAATTCGAGCTCCCTGTCCTGCTGAGAGACTCAGCAGCAGCAGCCTTAATCAAAG 2967  
Db 241 GAATTCGAGCTCCCTGTCCTGCTGAGAGACTCAGCAGCAGCAGCCTTAATCAAAG 300  
QY 2968 GTTTTACAGCCCTTTGGACACTATATGGGAGGAGGCAAGAGTACACCAATTTGTTAAAGC 3027  
Db 301 GTTTTACAGCCCTTTGGACACTATATGGGAGGAGGCAAGAGTACACCAATTTGTTAAAGC 360  
QY 3028 AAGAAACCAAGTCTCTTCTCACTAGTCAATTTAGACATGTTATCATCCAGACTATC 3087  
Db 361 AAGAAACCAAGTCTCTTCTCACTAGTCAATTTAGACATGTTATCATCCAGACTATC 420  
QY 3088 TACCCTGCAACATTCGAATCTCCCAAGAGCAAAATCCCACTTCTTGAAGTCTTGCAAGCTTC 3147  
Db 421 TACCCTGCAACATTCGAATCTCCCAAGAGCAAAATCCCACTTCTTGAAGTCTTGCAAGCTTC 480  
QY 3148 TGTGTAATATAGGCGAGCTGTCTGTATGCGGTAGAAATCAATGATCTGAGGACCAATTCAT 3207  
Db 481 TGTGTAATATAGGCGAGCTGTCTGTATGCGGTAGAAATCAATGATCTGAGGACCAATTCAT 540  
QY 3208 GGAAGCTGCTAAATAGCCTTAGTCTCGGGAGTCTTCCATAAAGTCTTGTGATGGAGCAACA 3267  
Db 541 GGAAGCTGCTAAATAGCCTTAGTCTCGGGAGTCTTCCATAAAGTCTTGTGATGGAGCAACA 600  
QY 3268 AACAGGATTAAACTAGGTTTGGTTCCTTACGCCCTCTCTAAAGCATAGGGCTTAGCCCTGCA 3327

601	AACAGGATTAAACTAGGTTTGGTTCTTCAGCCCTCTAAAGCATAGGGCTTAGCCTGCA	660
3328	GGCTTCCTTGGGGTTTCTCTGTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGAT	3387
661	GGCTTCCTTGGGGTTTCTCTGTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGAT	720
3388	CCAGTGTCATGGAACCTTCCACATGCCGTGACTCTGGACTATATCAGTTTTTGGAAA	3447
721	CCAGTGTCATGGAACCTTCCACATGCCGTGACTCTGGACTATATCAGTTTTTGGAAA	780
3448	GCAGGGTTCTCTGCGCTGTAAACAAGCCACGTCGAGCCAGTCGTGAATGCTTTCTCTTAC	3507
781	GCAGGGTTCTCTGCGCTGTAAACAAGCCACGTCGAGCCAGTCGTGAATGCTTTCTCTTAC	840
3508	ACCTATGTTTTTAAAGTAGTCAAACTTCAAGAAACAATCTAAACAAGTTTCTGTTCG	3563
841	ACCTATGTTTTTAAAGTAGTCAAACTTCAAGAAACAATCTAAACAAGTTTCTGTGTGC	896

```

RESULT 3
15-09-543-597-209
: Sequence 209, Application US/09643597
: Patent No. 6426072
: GENERAL INFORMATION:
: APPLICANT: Wang, Tongtong
: APPLICANT: Fan, Liguon

```

```

/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Baqur, Chaitanya S.
/ APPLICANT: Hosken, Nancy
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skeiky, Yasir A.W.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: McNeill, Patricia D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ OF LUNG CANCER
/ TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.455C11
/ CURRENT APPLICATION NUMBER: US/09/643,597
/ CURRENT FILING DATE: 2000-08-21
/ NUMBER OF SEQ ID NOS: 369
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 209

```

```

; LENGTH: 621
; TYPE: DNA
; ORGANISM: Homo sapiens
JS-09-643-597-209

```

Query Match 16.8%; Score 619.4; DB 4; Length 621;  
Best Local Similarity 99.8%; Pred. No. 8.4e-166;  
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

3055 CATTAGAACATGGTTATCATCCAGACTACTCTACCCTCGCAACATTGAACCTCCCAAGAG 3114

db 1 CATTTAGAACATGGTTATCATCCAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAG 60

3115 CAAATCCACATTCCTCTTGAGTTCTGCAGCTTCTGTGTAATAGGCAGCTGTCGTCTAT 3174

Db 61 CAAATCCACATTCCTCTTGTAGTTCTGCAGCTTCTGTGTAATAGGCAGCTGTCGTCTAT 120

Qy 3175 GCCGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCCTAGTCTGGG 3234

Db 121 GCCGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCCTAGTCTGG 180

Qy 3235 GAGTCTTCCATAAAGTTTTCATGGAGCAAAACAACAGGATTAAC TAGGTTGGTTGCT 3294

D**b** 181 GAGTCTTCCAATAAGTTTGCATGGAGCAACAACAGGATTAACACTAGTTTGGTTCCT 240

3295 TCAGCCCTCTAAAGCATAGGGCTTAGCCCTCAGCCTTCCTTGGGTTTCTCTGTGTGTG 3354

Db 241 TCAGCCCTCTAAAGCATAGGGCTTAGCCCTGAGGTTCTCTGGGTTTCTCTGTGTGTG 300

Qy	3355	TAGTTTGTAAACACATATAGCATCTGTTAAGATCCAGTCCAGTCCAGTAAACCTTCCACAT	3411
Db	301	TAGTTTGTAAACACATATAGCATCTGTTAAGATCCAGTCCAGTAAACCTTCCACAT	360
Qy	3415	GCCTGACTCTGCAGTACTATACAGTTTTTGGAAAGCAGGGTTCTCTGCCTGTAAACAAGC	3474
Db	361	GCCTGACTCTGCAGTACTATACAGTTTTTGGAAAGCAGGGTTCTCTGCCTGTAAACAAGC	420
Qy	3475	CCAGTGGACCAAGTCGAAATGCTTTCCCTTTACACCTATGTTTTAAAGTAGTCAAACTTC	3534
Db	421	CCAGTGGACCAAGTCGAAATGCTTTCCCTTTACACCTATGTTTTAAAGTAGTCAAACTTC	480
Qy	3535	AAGAAACAATCTAAACAAGTTCTGTGTGCATATGTTGTTGTGAACCTGTATTTGTATTTA	3594
Db	481	AAGAAACAATCTAAACAAGTTCTGTGTGCATATGTTGTTGTGAACCTGTATTTGTATTTA	540
Qy	3595	GTAGGCTTCATATTCGATTTAACTTGTTTTTCTGAATCTCTGATTCCTCTTTTCGGATA	3654
Db	541	GTAGGCTTCATATTCGATTTAACTTGTTTTTCTGAATCTCTGATTCCTCTTTTCGGATA	600
Qy	3655	CTATTGATGAATTAAGAAATT	3675
Db	601	CTATTGATGAATTAAGAAATT	621

```

RESULT 4
US-09-480-884A-209
; Sequence 209, Application US/09480884A
; Patent NO. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121-455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-480-884A-209

```

Query Match 16.8%; Score 619.4; DB 4; Length 621;  
Best Local Similarity 99.8%; Pred. No. 8.4e-166;  
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

3055 CATTAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAG 3114

1 CATTTAGAACATGGTTATCATCCAAGACTACTTACCCCTGCAACATTGAATCCCAAGAG 60

QY 3115 CAAATGCACATTCTCTTGAGTTCTGCAGCTTCGTGTGTAATAGGCGAGCTGTCGTCTAT 3174

Db 61 CAAATCCACATTCCTCTTGAGTTCTGCAGCTTGTGTAAATAGGCAGCTGTCGTCTAT 120

3175 GCCGTAGAA TCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCCTAGTCTGGG 3234

Db 121 GCCGTAGAATCACATGATCTGAGGACCATTATGGAAGCTGCTAATAGCCTAGTCTGGG 180

3235 GAGTCCTTCCATAAAGTTTTGCATGGAGCAAAACAACGATTAAC TAGTTTGGTTCT 3294

Db 181 GAGTCCTCCATAAAAGTTTTCATGGAGCAAAACAACAGGATTAACTAGTTTGGTTCTT 240

3295 TCAGCCCTCTAAAAGCATAGGGCTTAGCCCTGCAGACCTTCCTGGGCTTCTCTCTGTGTGTG 3355

Db 241 TCAGCCCTCTAAAGCATAGGCGCTTAGCCTGTCAGCGTCCTTGGCGCTTCTCTGCGTGGTG 300

QY 3355 TAGTTTGTATACACTATAGCAATCTGTTAAGAATCCAGAGTCCATGGGAAATCCATCCACAT 3361

Db 301 TAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGCCATGGAACCTTCCACAT 360  
2Y 3415 GCCGTGACTCTGGACTATATCAGTCTTTTGGAAAGCAGGGTTCCTCTGCTGCTTAACAAGC 3474  
Db 361 GCCGTGACTCTGGACTATATCAGTCTTTTGGAAAGCAGGGTTCCTCTGCTGCTTAACAAGC 420  
2Y 3475 CCAGTGGACCAAGTCTGAATGCTCTTCCCTTTTACACCTATGTTTTTAAGTAGTCAAACTTC 3534  
Db 421 CCAGTGGACCAAGTCTGAATGCTCTTCCCTTTTACACCTATGTTTTTAAGTAGTCAAACTTC 480  
QY 3535 AAGAAACAATCTAAACAAGTTCTCTGTGTCATATGTTTGTGAACCTTGTTATTTGTTATTTA 3594  
Db 481 AAGAAACAATCTAAACAAGTTCTCTGTGTCATATGTTTGTGAACCTTGTTATTTGTTATTTA 540  
QY 3595 GTAGCTTCTATATGCAATTTAACTTGTTTGTGAACCTTGTTATTTGTTATTTGTTATTTA 3654  
Db 541 GTAGCTTCTATATGCAATTTAACTTGTTTGTGAACCTTGTTATTTGTTATTTGTTATTTA 600  
QY 3655 CTATTGATGAATAAAGAAATT 3675  
Db 601 CTATTGATGAATAAAGAAATT 621

RESULT 5  
US-09-542-615A-209  
; Sequence 209, Application US/09542615A  
; Patent No. 6518256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C8  
; CURRENT APPLICATION NUMBER: US/09/542,615A  
; CURRENT FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 209  
; LENGTH: 621  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-542-615A-209

Query Match 16.8%; Score 619.4; DB 4; Length 621;  
Best Local Similarity 99.8%; Pred. No. 8.4e-166;  
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3055 CATTAGAACATGGTTATCATCAAGACTACTCTACCTGCAACATTGAACCTCCCAAGAG 3114  
Db 1 CATTAGAACATGGTTATCATCAAGACTACTCTACCTGCAACATTGAACCTCCCAAGAG 60  
QY 3115 CAAATCCACATTCCTCTGAGTTCTGACGTTCTGTGTAATAGGGCAGCTGTCGTCTAT 3174  
Db 61 CAAATCCACATTCCTCTGAGTTCTGACGTTCTGTGTAATAGGGCAGCTGTCGTCTAT 120  
QY 3175 GCCGTAGAATCAGATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCCCTAGTCTGG 3234  
Db 121 GCCGTAGAATCAGATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCCCTAGTCTGG 180  
QY 3235 GAGTCTTCCATAAAGTTTTCATGGAGCAACAAACAGGATTAACCTAGTTTGGTTCCT 3294  
Db 181 GAGTCTTCCATAAAGTTTTCATGGAGCAACAAACAGGATTAACCTAGTTTGGTTCCT 240  
QY 3295 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGACGGCTTCCCTGGGCTTCTCTGTGTGTG 3354  
Db 241 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGACGGCTTCCCTGGGCTTCTCTGTGTGTG 300  
QY 3355 TAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGCCATGGAACCTTCCACAT 3414

Db 301 TAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGCCATGGAACCTTCCACAT 360  
QY 3415 GCCGTGACTCTGGACTATATCAGTCTTTTGGAAAGCAGGGTTCCTCTGCTGCTTAACAAGC 3474  
Db 361 GCCGTGACTCTGGACTATATCAGTCTTTTGGAAAGCAGGGTTCCTCTGCTGCTTAACAAGC 420  
QY 3475 CCAGTGGACCAAGTCTGAATGCTCTTCCCTTTTACACCTATGTTTTTAAGTAGTCAAACTTC 3534  
Db 421 CCAGTGGACCAAGTCTGAATGCTCTTCCCTTTTACACCTATGTTTTTAAGTAGTCAAACTTC 480  
QY 3535 AAGAAACAATCTAAACAAGTTCTCTGTGTCATATGTTTGTGAACCTTGTTATTTGTTATTTA 3594  
Db 481 AAGAAACAATCTAAACAAGTTCTCTGTGTCATATGTTTGTGAACCTTGTTATTTGTTATTTA 540  
QY 3595 GTAGCTTCTATATGCAATTTAACTTGTTTGTGAACCTTGTTATTTGTTATTTGTTATTTA 3654  
Db 541 GTAGCTTCTATATGCAATTTAACTTGTTTGTGAACCTTGTTATTTGTTATTTGTTATTTA 600  
QY 3655 CTATTGATGAATAAAGAAATT 3675  
Db 601 CTATTGATGAATAAAGAAATT 621

RESULT 6  
US-09-606-421B-209  
; Sequence 209, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C9  
; CURRENT APPLICATION NUMBER: US/09/606,421B  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 209  
; LENGTH: 621  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-606-421B-209

Query Match 16.8%; Score 619.4; DB 4; Length 621;  
Best Local Similarity 99.8%; Pred. No. 8.4e-166;  
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3055 CATTAGAACATGGTTATCATCAAGACTACTCTACCTGCAACATTGAACCTCCCAAGAG 3114  
Db 1 CATTAGAACATGGTTATCATCAAGACTACTCTACCTGCAACATTGAACCTCCCAAGAG 60  
QY 3115 CAAATCCACATTCCTCTGAGTTCTGACGTTCTGTGTAATAGGGCAGCTGTCGTCTAT 3174  
Db 61 CAAATCCACATTCCTCTGAGTTCTGACGTTCTGTGTAATAGGGCAGCTGTCGTCTAT 120  
QY 3175 GCCGTAGAATCAGATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCCCTAGTCTGG 3234  
Db 121 GCCGTAGAATCAGATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCCCTAGTCTGG 180  
QY 3235 GAGTCTTCCATAAAGTTTTCATGGAGCAACAAACAGGATTAACCTAGTTTGGTTCCT 3294  
Db 181 GAGTCTTCCATAAAGTTTTCATGGAGCAACAAACAGGATTAACCTAGTTTGGTTCCT 240  
QY 3295 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGACGGCTTCCCTGGGCTTCTCTGTGTGTG 3354

Db 241 TCAGCCCTCTAAAAGCATAGGGCTTAGCCCTGAGGCTTCCTTGGGCTTTCTGTGTGTG 300  
QY 3355 TAGTTTGTAAACACTATAGCATCTGTAAAGATCCAGTGTCCATGAAACCTTCCACAT 3414  
Db 301 TAGTTTGTAAACACTATAGCATCTGTAAAGATCCAGTGTCCATGAAACCTTCCACAT 360  
QY 3415 GCGGTGACTCTGGACTATATCAAGTTTGTGGAAAGCAGGGTTCTCTGCTGCTAAACAGC 3474  
Db 361 GCGGTGACTCTGGACTATATCAAGTTTGTGGAAAGCAGGGTTCTCTGCTGCTAAACAGC 420  
QY 3475 CCAGGTGGACCACTGTAATGCTTCTTCCCTTACACTATGTTTAAAGTAGTCAAACTTC 3534  
Db 421 CCAGGTGGACCACTGTAATGCTTCTTCCCTTACACTATGTTTAAAGTAGTCAAACTTC 480  
QY 3535 AAGAAACAATCTAAACAAGTTTCTGTGTGATATGTTTGTGAACCTTGTATTTGATTTA 3594  
Db 481 AAGAAACAATCTAAACAAGTTTCTGTGTGATATGTTTGTGAACCTTGTATTTGATTTA 540  
QY 3595 GTAGGCTTCTATATTCATTTAACTTGTGTTTGTAACTTCTCTCTTTCGGATA 3654  
Db 541 GTAGGCTTCTATATTCATTTAACTTGTGTTTGTAACTTCTCTCTTTCGGATA 600  
QY 3655 CTATTGATGAATAAGAAATT 3675  
Db 601 CTATTGATGAATAAGAAATT 621

## RESULT 7

US-08-896-164-34/c

; Sequence 34, Application US/08896164

; Patent No. 6218521

; GENERAL INFORMATION:

; APPLICANT: OBATA, Yuichi

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED

; TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR

; TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER

; NUMBER OF SEQUENCES: 87

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felle &amp; Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/896,164

; FILING DATE: July 17, 1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6218521man D. Hanson

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 34:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 708 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: Double

; TOPOLOGY: linear

US-08-896-164-34

Query Match

12.9% Score 477.2; DB 3; Length 708;

Best Local Similarity

80.8%; Pred. No. 2.4e-125;

Matches 573; Conservative

0; Mismatches 131; Indels 5; Gaps 5;

QY

2374 CCAGCCCTTGACACTATGGAGGCGGCAAGAGTACACC-AATTGTTAAAGCAAGAA 3032

Db 706 CCAGCCCTTGAAAAATAAGGGGGGGGCAAGAAACACCNATTTGTTAAAAAGCAAGA 647  
QY 3033 ACCACAGTGTCTCTTCACTAGTCATTTAGAAACATGTTATCATCCAAAGACTACTCTACCC 3092  
Db 646 AACACAGGTTTTTTCATAGTATGTTAGAAAANGGTTTNNATCCAAGA-TAATTNACCC 588  
QY 3093 TGAACATTGAATCCCAAGAGCAAAATCCACATTCCTCTTGAGTTCTTGAGCTTCTGTGT 3152  
Db 587 GGNAAACATTG-ANTCCNAGNGCNAATTCACATTCCTGAGTTCCTGAGGTTTGCAGGTTCCGNGT 529  
QY 3153 AAATAGGGCAGCTGTCTGTCTATCCCGTAGAATCACATGATCTGAGGACCATTCATCGAAG 3212  
Db 528 AAATAGGGCAGTGTGTTTATGCCGTAGNATNANATGATTTGGGNGCCATTNATGAGNAG 469  
QY 3213 CTGCTAAATAGCCTAGTCTGGGAGTCTTCCATAAAGTTTTCATGAGGCAAAACAACAG 3272  
Db 468 NNGNTAAATAGCCTAGTCTGGGG-GTTTTCATATAAAGTTTTCATGAGGCAAAACAACAG 410  
QY 3273 GATTAAACAGGTTTGGTTCCCTTCAGCCCTCTATAAAGCATAGGCTTAGCCTGCAGGCTT 3332  
Db 409 GNTTAAATAGGTTTGGTTCCCTTCAGCCCTTATAAAGNATAGGNTTAGCCTGCAGG-TT 351  
QY 3333 CCTTGGGCTTCTCTGTGTGTGTGTTTGTGAAACACTATAGCATCTGTTAAGATCCAGT 3392  
Db 350 CCTTGGGCTTCTTGTGNGGNGTAGTTTGTAAANANTATAGCATTTGTTAAGATNCAGN 291  
QY 3393 GTCATGGAACCTTCCCATGCTCCGCTGACTCTGGAATATATCAGTTTGTGAAAGCAGG 3452  
Db 290 GTCATGGAACCTTCCCATGCTCCGCTGACTCTGGAATATATCAGTTTGTGAAAGCAGG 231  
QY 3453 GTTCTCTGCTCTAAACAGCCAGCTGGAACAGTCTGAATGCTTCTTCTTACACCTA 3512  
Db 230 GTTCTTCTGCTCTGTAANAAGCCCAAGTGTGACAGTTTGNANGTTTCTTCTTANACCTA 171  
QY 3513 TGTTTTAAGTAGTCAAACTTCAAGAAACATCTAAACAGTTTCTGTTGCATATGTT 3572  
Db 170 TGNTTTAAAGTAGTNNAAATTAAGGAACNAITTAACCAAGTTTGTGTCATANGTT 111  
QY 3573 TGTGAACCTTGTATTTGTTATTAGTAGGCTTCTATATTGCAATTAACCTGTTTGTGTAAC 3632  
Db 110 TGTGNAATTTGTTATTGTTATTAGTAGGTTTATATTGCAATTAACCTGTTTGTGTAAC 51  
QY 3633 CTGATTTCTTCTTCTTCCGATCTATTGATGAATAAAGAAATTAAGTG 3681  
Db 50 CCTGATTTTCTTCTTNGGATAATATTGTTGNATAAAGGAATTAAGTG 2

## RESULT 8

US-09-643-597-315

; Sequence 315, Application US/09643597

; Patent No. 6426072

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Henderson, Robert A.

; APPLICANT: McNeill, Patricia D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C11

; CURRENT APPLICATION NUMBER: US/09/643,597

; CURRENT FILING DATE: 2000-08-21

; NUMBER OF SEQ ID NOS: 369

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 315

; LENGTH: 336

TYPE: DNA  
ORGANISM: Homo sapien  
US-09-480-884A-315  
Query Match 9.1%; Score 336; DB 4; Length 336;  
Best Local Similarity 100.0%; Pred. No. 2e-85;  
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3057 TTGAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 3116  
DB 1 TTGAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 60  
QY 3117 AATCCACATTCCTCTTGAGTTCTGAGCTTCTGAGTAAATAGGCGAGCTGTCGTCTATGC 3176  
DB 61 AATCCACATTCCTCTTGAGTTCTGAGCTTCTGAGTAAATAGGCGAGCTGTCGTCTATGC 120  
QY 3177 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 3236  
DB 121 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 180  
QY 3237 GTCTTCCATAAAGTTTTCATGGAGCAAAACAAACAGGATTAAATAGGTTTGGTTCTTTC 3296  
DB 181 GTCTTCCATAAAGTTTTCATGGAGCAAAACAAACAGGATTAAATAGGTTTGGTTCTTTC 240  
QY 3297 AGCCCTCTAAAGCATAGGGCTTAGCCCTGCAAGCTTCTTGGGCTTCTCTGTGTGTGTA 3356  
DB 241 AGCCCTCTAAAGCATAGGGCTTAGCCCTGCAAGCTTCTTGGGCTTCTCTGTGTGTGTA 300  
QY 3357 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 3392  
DB 301 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 336  
RESULT 9  
US-09-480-884A-315  
; Sequence 315, Application US/09480884A  
; Patent No. 6482597  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C6  
; CURRENT APPLICATION NUMBER: US/09/480,884A  
; NUMBER OF SEQ ID NOS: 330  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 315  
; LENGTH: 336  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-480-884A-315

Query Match 9.1%; Score 336; DB 4; Length 336;  
Best Local Similarity 100.0%; Pred. No. 2e-85;  
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3057 TTGAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 3116  
DB 1 TTGAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 60  
QY 3117 AATCCACATTCCTCTTGAGTTCTGAGCTTCTGAGTAAATAGGCGAGCTGTCGTCTATGC 3176  
DB 61 AATCCACATTCCTCTTGAGTTCTGAGCTTCTGAGTAAATAGGCGAGCTGTCGTCTATGC 120  
QY 3177 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 3236  
DB 121 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 180  
QY 3237 GTCTTCCATAAAGTTTTCATGGAGCAAAACAAACAGGATTAAATAGGTTTGGTTCTTTC 3296

DB 181 GTCTTCCATAAAGTTTTCATGGAGCAAAACAAACAGGATTAAATAGGTTTGGTTCTTTC 240  
QY 3297 AGCCCTCTAAAGCATAGGGCTTAGCCCTGCAAGCTTCTTGGGCTTCTCTGTGTGTGTA 3356  
DB 241 AGCCCTCTAAAGCATAGGGCTTAGCCCTGCAAGCTTCTTGGGCTTCTCTGTGTGTGTA 300  
QY 3357 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 3392  
DB 301 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 336  
RESULT 10  
US-09-542-615A-315  
; Sequence 315, Application US/09542615A  
; Patent No. 6518256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C8  
; CURRENT APPLICATION NUMBER: US/09/542,615A  
; CURRENT FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 315  
; LENGTH: 336  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-542-615A-315  
Query Match 9.1%; Score 336; DB 4; Length 336;  
Best Local Similarity 100.0%; Pred. No. 2e-85;  
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3057 TTGAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 3116  
DB 1 TTGAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 60  
QY 3117 AATCCACATTCCTCTTGAGTTCTGAGCTTCTGAGTAAATAGGCGAGCTGTCGTCTATGC 3176  
DB 61 AATCCACATTCCTCTTGAGTTCTGAGCTTCTGAGTAAATAGGCGAGCTGTCGTCTATGC 120  
QY 3177 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 3236  
DB 121 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 180  
QY 3237 GTCTTCCATAAAGTTTTCATGGAGCAAAACAAACAGGATTAAATAGGTTTGGTTCTTTC 3296  
DB 181 GTCTTCCATAAAGTTTTCATGGAGCAAAACAAACAGGATTAAATAGGTTTGGTTCTTTC 240  
QY 3297 AGCCCTCTAAAGCATAGGGCTTAGCCCTGCAAGCTTCTTGGGCTTCTCTGTGTGTGTA 3356  
DB 241 AGCCCTCTAAAGCATAGGGCTTAGCCCTGCAAGCTTCTTGGGCTTCTCTGTGTGTGTA 300  
QY 3357 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 3392  
DB 301 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 336

RESULT 11  
US-09-606-421B-315  
; Sequence 315, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
US-09-606-421B-315



```
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606.421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 315
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-315

Query Match          9.1%; Score 336; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 2e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3057 TTTAGAACATGGTTATATCATCCAGACTACTCTACCTGCAACATTGAATCCCAAGACCA 3116
DB 1 TTTAGAACATGGTTATATCATCCAGACTACTCTACCTGCAACATTGAATCCCAAGACCA 60

QY 3117 AATCCACATTCCTCTTGAGTTCTGCAGCTTCTGTGTAATAGGCGAGCTGTCGTCATGC 3176
DB 61 AATCCACATTCCTCTTGAGTTCTGCAGCTTCTGTGTAATAGGCGAGCTGTCGTCATGC 120

QY 3177 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGTCTGGGGA 3236
DB 121 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGTCTGGGGA 180

QY 3237 GCTTCCATTAAGTTTTCATGAGCAACAAACAGGATTAACTAGTTTGGTTCTTC 3296
DB 181 GCTTCCATTAAGTTTTCATGAGCAACAAACAGGATTAACTAGTTTGGTTCTTC 240

QY 3297 AGCCCTCTAAAAGCATAGGGCTTAGCCCTGACAGGCTTCCTGGGCTTCTCTGTGTGTA 3356
DB 241 AGCCCTCTAAAAGCATAGGGCTTAGCCCTGACAGGCTTCCTGGGCTTCTCTGTGTGTA 300

QY 3357 GTTTTGTAAACACTATAGCATCTGTTAGATCCAGT 3392
DB 301 GTTTTGTAAACACTATAGCATCTGTTAGATCCAGT 336
```

```
RESULT 12
US-09-412-554A-7
; Sequence 7, Application US/09412554A
; Patent No. 6355788
; GENERAL INFORMATION:
; APPLICANT: Konklin, Darrell
; APPLICANT: Ellsworth, Jeff L.
; TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN ZFSTA2
; FILE REFERENCE: 98-50
; CURRENT APPLICATION NUMBER: US/09/412,554A
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1255
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 3' end of zfsta2 nucleotide sequence
```

```
Query Match          1.9%; Score 70.2; DB 4; Length 1255;
Best Local Similarity 64.4%; Pred. No. 1.8e-09;
Matches 105; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
```

```
QY 214 CACAGAGAAAGGGGAAACCCACCTGCTCTCTGATGAGCAATGCAAACTCACAAGAGGCC 273
DB 294 CAGAGAGACAGGGCAAGCAGAGATGCGCTGTATGACCTTTGCAAACTCCTACTACAAACC 353
QY 274 TGTGTGTGCGCATATATGCGAAGACCTACTCAACCACTGTGAATCGCATCGAGATGCCCTG 333
DB 354 TGTGTGTGATCTGACGCGAGAAATCTATGAAAAACCACTGTGAAGTGCACAGAGCTGCTTG 413
QY 334 CTTCACTGGATCCAAAATCCAGGTTGATTAGATGGACACTGC 376
DB 414 CCTGAAAAAACAAAAGATTACCAATGTTTCAATGAGACTGC 456
```

```
RESULT 13
US-09-412-554A-1
; Sequence 1, Application US/09412554A
; Patent No. 6355788
; GENERAL INFORMATION:
; APPLICANT: Konklin, Darrell
; APPLICANT: Ellsworth, Jeff L.
; TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN ZFSTA2
; FILE REFERENCE: 98-50
; CURRENT APPLICATION NUMBER: US/09/412,554A
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3192
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58) ... (3006)
US-09-412-554A-1
```

```
Query Match          1.9%; Score 70.2; DB 4; Length 3192;
Best Local Similarity 64.4%; Pred. No. 3.3e-09;
Matches 105; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
```

```
QY 214 CACAGAGAAAGGGGAAACCCACCTGCTCTCTGATGAGCAATGCAAACTCACAAGAGGCC 273
DB 294 CAGAGAGACAGGGCAAGCAGAGATGCGCTGTATGACCTTTGCAAACTCCTACTACAAACC 353
QY 274 TGTGTGTGCGCATATATGCGAAGACCTACTCAACCACTGTGAATCGCATCGAGATGCCCTG 333
DB 354 TGTGTGTGATCTGACGCGAGAAATCTATGAAAAACCACTGTGAAGTGCACAGAGCTGCTTG 413
QY 334 CTTCACTGGATCCAAAATCCAGGTTGATTAGATGGACACTGC 376
DB 414 CCTGAAAAAACAAAAGATTACCAATGTTTCAATGAGACTGC 456
```

```
RESULT 14
US-09-412-554A-3
; Sequence 3, Application US/09412554A
; Patent No. 6355788
; GENERAL INFORMATION:
; APPLICANT: Konklin, Darrell
; APPLICANT: Ellsworth, Jeff L.
; TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN ZFSTA2
; FILE REFERENCE: 98-50
; CURRENT APPLICATION NUMBER: US/09/412,554A
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2949
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate oligonucleotide sequence encoding the
; OTHER INFORMATION: zfsta2 polypeptide of SEQ ID NO:2.
; NAME/KEY: variation
```

LOCATION: (1)....(2949)  
OTHER INFORMATION: Each N is independently any nucleotide.  
US-09-412-554A-3

Query Match 1.8%; Score 66.8; DB 4; Length 2949;  
Best Local Similarity 41.5%; Pred. No. 2.9e-08;  
Matches 86; Conservative 42; Mismatches 76; Indels 3; Gaps 1;  
2y 173 TGTCCCAATGTGTTTGTGGAGCGCGCGGAATGTG--CAGTCACAGAGAAGGGGA 229  
Db 193 TGYGARAAYARTAYTGYGNYTNGMNGNCAYTGYTNACNMSNMNGNARNGNCAR 252  
2y 230 CCCACCTGTCTCTCAATGAGCAATGCAACCTCACAGAGGCGCTGTGTGTCAGTAAT 289  
Db 253 GCGAR\*GYGCGTGYATGAYTNTGYAARMGNCAYTAYAACCGNGTNTGYGWNNGAY 312  
Qy 290 GCGAAGACCTACCTCAACCTGTGACTGATGATGATGATGATGATGATGATGATGAT 349  
Db 313 GCGARTTYTAYGARAAYCAYTGYGAGTNGMNGNGCNGTNGYTTNARAARCAPAAR 372  
Qy 350 ATCCAGGTTGATTACGATGGACACTGC 376  
Db 373 ATHACNATHGTCNAYAXGARGATGY 399

RESULT 15  
US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ99c-Fls  
US-08-232-463-14

Query Match 1.4%; Score 53.2; DB 1; Length 7218;  
Best Local Similarity 4.5%; Pred. No. 0.00039;  
Matches 13; Conservative 172; Mismatches 105; Indels 0; Gaps 0;  
Qy 770 GAGAGAAGTGTGCTGAGAGATGAAACGATGATGAGATGAGATGAGATGAGATGAGAT 829  
Db 1331 RRR 1272  
Qy 830 TGTAAACCGTGTGTCTGTGCTGTGAAATTTGGTCTGTACAGCCATGACCTGTGACGGA 889  
Db 1271 RRR 1212  
Qy 890 AAGAATCAGAAAGGGGCGCCAGACAGAGAGAGAGATGACAGAGATATGTCCAGGAG 949  
Db 1211 RRR 1152  
Qy 950 CTCCAAAGCATCAGAAACAGCTGAAAGACCAAGAGAGTGAAGAGAGATCTAA 1009  
Db 1151 RRR 1092  
Qy 1010 TGAGGAGGACAGACAGTGTGTGATCCAGCATCTTCTTCCACTTCAGC 1059  
Db 1091 RRR 1042

Search completed: March 11, 2004, 14:41:33  
Job time : 172.644 secs



c

RESIST. 1

	Query Match	100.0%;	Score 3693;	DB 10;	Length 3693;	
	Best Local Similarity	100.0%;	Pred. No. 0;			
	Matches 3693;	Conservative	0;	Mismatches	0;	Indels
						Gaps
						0
QY	1	CGATCGCGGACTCCCACTCCGCTTACAGTCGCTGCGCGCGCGTCTGCCCGCGCGCCCC	60			
DB	1	CGATCGCGGACTCCCACTCCGCTTACAGTCGCTGCGCGCGTCTGCCCGCGCGCCCC	60			
QY	61	AGGAGACTCGACCGACACACGATGTGGAAA	CGCTGGCTCGGCTCGCGCTCGCGCTGTGT	120		
DB	61	AGGAGACTCGACCGACACGATGTGGAAA	CGCTGGCTCGGCTCGCGCTCGCGCTGTGT	120		
QY	121	GGCGGCTCGCTGGGTCGCGCCGAGGAAGCTAAGGAGCAAT	CCAAGATCTGTGCCAA	180		
DB	121	GGCGGCTCGCTGGGTCGCGCCGAGGAAGCTAAGGAGCAAT	CCAAGATCTGTGCCAA	180		

Database : Published Applications NA:\*

## SUMMARIES

Result No.	Score	Query #		DB	ID	Description
		Match	Length			
1	3693	100.0	3693	10	US-09-019-039-359	Sequence 359, App
2	3693	100.0	3693	14	US-10-084-817-20	Sequence 20, App1
3	1340	52.5	1987	9	US-09-830-107-3287	Sequence 3287, App
4	1340	52.5	1987	10	US-09-860-706-941	Sequence 941, App
5	1340	52.5	1987	10	US-09-873-319-617	Sequence 617, App
6	1729.8	46.8	2655	13	US-10-098-684-144	Sequence 144, App
7	1678.4	45.4	1755	9	US-09-754-877-3903	Sequence 3903, App
8	1678.4	45.4	1755	15	US-10-442-513-3903	Sequence 3903, App
9	630.4	17.1	634	9	US-09-738-973-540	Sequence 540, App
10	630.4	17.1	634	9	US-09-738-973-540	Sequence 540, App
11	630.4	17.1	634	14	US-10-854-133-540	Sequence 540, App
12	619.4	16.8	621	9	US-09-735-705-209	Sequence 209, App
13	619.4	16.8	621	9	US-09-850-176A-209	Sequence 209, App
14	619.4	16.8	621	9	US-09-837-778-209	Sequence 209, App
15	619.4	16.8	621	10	US-09-466-396A-209	Sequence 209, App

181 TGTGTTTTGTGAGCGCGCGGAAATGTGCACTCAAGAGAAAGGGAAACCCACCTGCT 240  
181 TGTGTTTTGTGAGCGCGCGGAAATGTGCACTCAAGAGAAAGGGAAACCCACCTGCT 240  
241 CTGCAATGAGCAATGCAAACTCTCAAGAGGCTGTGTGACGATTAATGCGAGACCTA 300  
241 CTGCAATGAGCAATGCAAACTCTCAAGAGGCTGTGTGACGATTAATGCGAGACCTA 300  
301 CCTCAACCACTGTGAATCTCATCGAGATGCTGCTCATCTGGATCCAAATCCAGGTTGA 360  
301 CCTCAACCACTGTGAATCTCATCGAGATGCTGCTCATCTGGATCCAAATCCAGGTTGA 360  
361 TTACGATGACACTGCAAGAGAGAAATCCCTAGTCCATCTGCGAGCCAGGTTGTTG 420  
361 TTACGATGACACTGCAAGAGAGAAATCCCTAGTCCATCTGCGAGCCAGGTTGTTG 420  
421 CTATCAGTCCAAACCGTGATGAGCTCCGAGTCCGATCATCTGCAAGTGGTGAAGCTGAGAT 480  
421 CTATCAGTCCAAACCGTGATGAGCTCCGAGTCCGATCATCTGCAAGTGGTGAAGCTGAGAT 480  
481 CATCTCAGATGGCTGTTCTTAAGGAGCAACTACAGTGAATCCCTAGACAAATATT 540  
481 CATCTCAGATGGCTGTTCTTAAGGAGCAACTACAGTGAATCCCTAGACAAATATT 540  
541 TAAGAACTTTGATAATGGTGATTTCTGCGCTGGACTCCAGTGAATTCCTGAAGTTGTGA 600  
541 TAAGAACTTTGATAATGGTGATTTCTGCGCTGGACTCCAGTGAATTCCTGAAGTTGTGA 600  
601 ACAGATGAACCTGCCATCAATATTAACAGTATCCAGACCGAGGAGAAACAAAGTTGCT 660  
601 ACAGATGAACCTGCCATCAATATTAACAGTATCCAGACCGAGGAGAAACAAAGTTGCT 660  
661 TAGGGAATCTGTGTTGATGCTCTCATGAACTGTCTGATGAATATGCTGATGAAACT 720  
661 TAGGGAATCTGTGTTGATGCTCTCATGAACTGTCTGATGAATATGCTGATGAAACT 720  
721 CAGCTTCAAGAGTTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCCTGAGAAGAGTG 780  
721 CAGCTTCAAGAGTTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCCTGAGAAGAGTG 780  
781 TGCCCTGAGGATGAACGATATGAGATGAGCTGAGACGAGTGGATGCTGAACCGCTG 840  
781 TGCCCTGAGGATGAACGATATGAGATGAGCTGAGACGAGTGGATGCTGAACCGCTG 840  
841 TGTCTGCTGCTGGAATTTGGTCTGTACAGCATGACCTGTGACGGAAGAAATCAGAA 900  
841 TGTCTGCTGCTGGAATTTGGTCTGTACAGCATGACCTGTGACGGAAGAAATCAGAA 900  
901 GGGGCCCCAGACCCAGACAGAGGAGATGACCATATGTCACGAGCTCCAAAGCA 960  
901 GGGGCCCCAGACCCAGACAGAGGAGATGACCATATGTCACGAGCTCCAAAGCA 960  
961 TCAGGAACAGCTGAAAGACCAAGAGAGTGAGCACCAAGAGATCTAATGAGGAGGCAC 1020  
961 TCAGGAACAGCTGAAAGACCAAGAGAGTGAGCACCAAGAGATCTAATGAGGAGGCAC 1020  
1021 AGACCACTGTGATCCAGCATCTTCTCCATCTCAGCGCTGAGTTCAATACACAG 1080  
1021 AGACCACTGTGATCCAGCATCTTCTCCATCTCAGCGCTGAGTTCAATACACAG 1080  
1081 TGTCTGTACAGTCCCAATCACCAGTATTTGCTTATATAGCAATGAGTTTATTTGT 1140  
1081 TGTCTGTACAGTCCCAATCACCAGTATTTGCTTATATAGCAATGAGTTTATTTGT 1140  
1141 TTATTTGTTTGCATTAAGGATATGAGTGGCTGCTAGGAGGAGGAGGCGCACAGC 1200  
1141 TTATTTGTTTGCATTAAGGATATGAGTGGCTGCTAGGAGGAGGAGGCGCACAGC 1200  
1201 TTCAATTTCTAGAGTGTCTTTAAGAGAACTGTAAATGGTGTCTGTGGGCTGAGGCTAGT 1260  
1201 TTCAATTTCTAGAGTGTCTTTAAGAGAACTGTAAATGGTGTCTGTGGGCTGAGGCTAGT 1260

1261 AAGAAACTGCATCAGATTGAAAGAGGAAACAGACCCAAATCTGACCTCTTTTGAGTTT 1320  
1261 AAGAAACTGCATCAGATTGAAAGAGGAAACAGACCCAAATCTGAACTCTTTTGAGTTT 1320  
1321 ACTGCATCTGTCAAGCAGGCTGCAAGGAGTGCAACGATCCAGAGAGAACTTAGCAGGTT 1380  
1321 ACTGCATCTGTCAAGCAGGCTGCAAGGAGTGCAACGATCCAGAGAGAACTTAGCAGGTT 1380  
1381 GTCCCGGAGGAGAGGTTTGGGAGCTCCACGAGAGGAAACGCTCTCTGCTCCAGCCTC 1440  
1381 GTCCCGGAGGAGAGGTTTGGGAGCTCCACGAGAGGAAACGCTCTCTGCTCCAGCCTC 1440  
1441 TTTCCATTGCGCTCAGCATGACAGACTCCAGCATCCAGCATCTCTTGGTCCCAATAAC 1500  
1441 TTTCCATTGCGCTCAGCATGACAGACTCCAGCATCCAGCATCTCTTGGTCCCAATAAC 1500  
1501 TGCTCTAGATATATAGCCATATCTGTAGTTAACCCAGTGTCCCTCAGACTTGGATGGAG 1560  
1501 TGCTCTAGATATATAGCCATATCTGTAGTTAACCCAGTGTCCCTCAGACTTGGATGGAG 1560  
1561 TTTCTGGAGGAGTACACCCAAATGATGCAGATCTTGTATACCTTGGAGCCCTTAGCGAC 1620  
1561 TTTCTGGAGGAGTACACCCAAATGATGCAGATCTTGTATACCTTGGAGCCCTTAGCGAC 1620  
1621 CTAAACCAATTTTAAATACTTTTACAAAGGCTGATTTCTCTGTAACACACTTTT 1680  
1621 CTAAACCAATTTTAAATACTTTTACAAAGGCTGATTTCTCTGTAACACACTTTT 1680  
1681 TTTGGCAAGTGTACTTATTTCAATTATTAATATATATATATATATATATATATAT 1740  
1681 TTTGGCAAGTGTACTTATTTCAATTATTAATATATATATATATATATATATATAT 1740  
1741 ATTTTCTGACTAGTATTAAGCTTTTGTAAATATTTTTCAGTAGTCCCAGCTTCTATA 1800  
1741 ATTTTCTGACTAGTATTAAGCTTTTGTAAATATTTTTCAGTAGTCCCAGCTTCTATA 1800  
1801 GGTGGAAGGAGTTTGGGTTCTTCTGCTGAGGGCTGAAATAACCCAGATGCTCCAC 1860  
1801 GGTGGAAGGAGTTTGGGTTCTTCTGCTGAGGGCTGAAATAACCCAGATGCTCCAC 1860  
1861 CCTGCCACATCTAGATGCAAGCCATAGTTGGGCCCTAGCTTCCAGAGTCCACTATC 1920  
1861 CCTGCCACATCTAGATGCAAGCCATAGTTGGGCCCTAGCTTCCAGAGTCCACTATC 1920  
1921 TGCCAGAGGAGCAAGGCTGCTTAGACTAAGCCAGGGGAGAGCATCTTCAATAAAAA 1980  
1921 TGCCAGAGGAGCAAGGCTGCTTAGACTAAGCCAGGGGAGAGCATCTTCAATAAAAA 1980  
1981 CTTTCAAGATCCAAACATTAATTTTATTTTATTTATTTCTGAGAGTTGAGGCAATCAGT 2040  
1981 CTTTCAAGATCCAAACATTAATTTTATTTTATTTATTTCTGAGAGTTGAGGCAATCAGT 2040  
2041 ATTTCCAGGATGGCAACAGGCGAGCCAGGCTTAGGATATCCAGCTTACCAAT 2100  
2041 ATTTCCAGGATGGCAACAGGCGAGCCAGGCTTAGGATATCCAGCTTACCAAT 2100  
2101 ATGCTCATTTCCACTAACTAGGAGGAGTTGGCCCTGTCTCTTTTCTTGGAGCTC 2160  
2101 ATGCTCATTTCCACTAACTAGGAGGAGTTGGCCCTGTCTCTTTTCTTGGAGCTC 2160  
2161 AGTTTCTCAGTGTGTTGAGTAATGCACTAACTTTTGTATTTGATTAAGTTATAATTC 2220  
2161 AGTTTCTCAGTGTGTTGAGTAATGCACTAACTTTTGTATTTGATTAAGTTATAATTC 2220  
2221 TGTGTTCTGATCATTTGGTCCAGAGGAGATAGTTTCTTGTGATTTTCTTCTCTCT 2280  
2221 TGTGTTCTGATCATTTGGTCCAGAGGAGATAGTTTCTTGTGATTTTCTTCTCTCT 2280  
2281 ATAGATAAATGAATCTTGTACTAGAAACAGAAATGTCAGATGGCCAAAAACAAGTG 2340  
2281 ATAGATAAATGAATCTTGTACTAGAAACAGAAATGTCAGATGGCCAAAAACAAGTG 2340  
2341 ACCAGATTGATCTCAGCCTGATGACCCCTACAGGCTGCTGATGATGAGTCTCTCATG 2400





QY	361	TTACGATCGACACCTGCAAGAGAGAAATCCGTAAGTCCATCTGCCAGCCAGTGTGTTG	420
Db	361	TTACGATCGACACCTGCAAGAGAGAAATCCGTAAGTCCATCTGCCAGCCAGTGTGTTG	420
QY	421	CTATCAGTCCCAACCGTGTAGTCCGACGTCGATCATCATCAGTGGCTGGAAGCTGAGAT	480
Db	421	CTATCAGTCCCAACCGTGTAGTCCGACGTCGATCATCATCAGTGGCTGGAAGCTGAGAT	480
QY	481	CATCTCAGATGGCTGTTCTCTAAGGAGCAACTACAGTGAATCTCAGCAAGTATTTT	540
Db	481	CATCTCAGATGGCTGTTCTCTAAGGAGCAACTACAGTGAATCTCAGCAAGTATTTT	540
QY	541	TAAGAACTTTGATATGTTGATCTCTCCCTGGACTCCAGTGAATCTCAGTGTGTTGGA	600
Db	541	TAAGAACTTTGATATGTTGATCTCTCCCTGGACTCCAGTGAATCTCAGTGTGTTGGA	600
QY	601	ACAGAACTGAACTCCGATCAATATTAACAAGTATCCAGAGCAGAGAAACAAGTGTCT	660
Db	601	ACAGAACTGAACTCCGATCAATATTAACAAGTATCCAGAGCAGAGAAACAAGTGTCT	660
QY	661	TAGGGACTCTGTGTGATGCTCTCATTTGAATCTCTGATGAATCTGATTTGGAACCT	720
Db	661	TAGGGACTCTGTGTGATGCTCTCATTTGAATCTCTGATGAATCTGATTTGGAACCT	720
QY	721	CAGCTTCCAGAGTCTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCTCCTGAGAAGATG	780
Db	721	CAGCTTCCAGAGTCTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCTCCTGAGAAGATG	780
QY	781	TGCCCTGAGGATGAACGATGAGATGAGAGTGCAGACGAGTGCATGTAAACGCTG	840
Db	781	TGCCCTGAGGATGAACGATGAGATGAGAGTGCAGACGAGTGCATGTAAACGCTG	840
QY	841	TGTCTGCTGTGAAATTTGGGTCTGTACAGCCATGACCTGTGACGGAAGAAATCAGAA	900
Db	841	TGTCTGCTGTGAAATTTGGGTCTGTACAGCCATGACCTGTGACGGAAGAAATCAGAA	900
QY	901	GGGGCCCCAGACCCAGACAGAGAGGAGATGACACAGATATGTCCAGAGCTCCAAAGCA	960
Db	901	GGGGCCCCAGACCCAGACAGAGAGGAGATGACACAGATATGTCCAGAGCTCCAAAGCA	960
QY	961	TCAGAAACAGCTGAAAGACCAAGAGAGTGCAGACCAAGAGATCTAATGAGAGGCAC	1020
Db	961	TCAGAAACAGCTGAAAGACCAAGAGAGTGCAGACCAAGAGATCTAATGAGAGGCAC	1020
QY	1021	AGACAGTGTCTGGATCCAGCATCTTCTCCATCTCAGCGCTGAGTTCAGTATACACAAG	1080
Db	1021	AGACAGTGTCTGGATCCAGCATCTTCTCCATCTCAGCGCTGAGTTCAGTATACACAAG	1080
QY	1081	TGTCTGTACAGTCCCAATTCACAGTATTTGCTTATATAGCAATCAGTATTTTGT	1140
Db	1081	TGTCTGTACAGTCCCAATTCACAGTATTTGCTTATATAGCAATCAGTATTTTGT	1140
QY	1141	TTATTTGTTTTCGAATAAGGATATGAAGTGGCTGAGGAGGAAAGGCGCACAGCC	1200
Db	1141	TTATTTGTTTTCGAATAAGGATATGAAGTGGCTGAGGAGGAAAGGCGCACAGCC	1200
QY	1201	TTCAATTTCTAGAGTGTCTTAAGAGAACTCTAATGGTGTCTGGGCTGGAGGCTAGT	1260
Db	1201	TTCAATTTCTAGAGTGTCTTAAGAGAACTCTAATGGTGTCTGGGCTGGAGGCTAGT	1260
QY	1261	AAGGAACTGTCATCAGATTTGAAAGAGGAAACAGACCCAAATCTGAACTCTTTTGTGTTT	1320
Db	1261	AAGGAACTGTCATCAGATTTGAAAGAGGAAACAGACCCAAATCTGAACTCTTTTGTGTTT	1320
QY	1321	ACTGATCTGTCCAGAGCTGCGAGGAGTGCACAGATGCCAGAGAGAACTTAGCAGGGT	1380
Db	1321	ACTGATCTGTCCAGAGCTGCGAGGAGTGCACAGATGCCAGAGAGAACTTAGCAGGGT	1380
QY	1381	GTCCCGGAGGAGAGGTTTGGGAAGCTCCAGGAGAGGAACGCTCTCTGTTTCCAGGCTC	1440
Db	1381	GTCCCGGAGGAGAGGTTTGGGAAGCTCCAGGAGAGGAACGCTCTCTGTTTCCAGGCTC	1440
QY	1441	TTTCCATTTGCCCTGACGATGACAGACCTTCAGCATCCACGCAATCTCTTGTGTTCCCAATAAC	1500

Db	1441	TTTCCATTTGCCCTGACGATGACAGACCTTCAGCATCCACGCAATCTCTTGTGTTCCCAATAAC	1500
QY	1501	TGCCTCTAGATACATAGCCATACCTAGTTAAACCCAGTGTCCCTCAGACTTGGATGGAG	1560
Db	1501	TGCCTCTAGATACATAGCCATACCTAGTTAAACCCAGTGTCCCTCAGACTTGGATGGAG	1560
QY	1561	TTTCTGGAGGGGTACCCCAATGATGAGATACCTGTATATCTTTGAGCCCTTAGCCAC	1620
Db	1561	TTTCTGGAGGGGTACCCCAATGATGAGATACCTGTATATCTTTGAGCCCTTAGCCAC	1620
QY	1621	CTAACCAAAATTTTAAAAATACCTTTTACAAAGGTGCTATTCTCTGTAAACACACTTTT	1680
Db	1621	CTAACCAAAATTTTAAAAATACCTTTTACAAAGGTGCTATTCTCTGTAAACACACTTTT	1680
QY	1681	TTTGGCAAGTTGACTTTTATCTTCAATTTATTTATCATATATTTGTTTAAATATTTT	1740
Db	1681	TTTGGCAAGTTGACTTTTATCTTCAATTTATTTATCATATATTTGTTTAAATATTTT	1740
QY	1741	ATTTTCTTGACTAGGTATTAAGCTTTTGTAAATTTTTCAGTAGTCCACCACTTCATA	1800
Db	1741	ATTTTCTTGACTAGGTATTAAGCTTTTGTAAATTTTTCAGTAGTCCACCACTTCATA	1800
QY	1801	GGTGAAGAGGTTTGGGTCTTCTGTGTGAGGGGCTGAAATAACCCAGATGTCTCCAC	1860
Db	1801	GGTGAAGAGGTTTGGGTCTTCTGTGTGAGGGGCTGAAATAACCCAGATGTCTCCAC	1860
QY	1861	CCTGCCACATATAGATGAGCCCATAGTTGGCCCCCTAGCTTCCAGCAGTCCACTATC	1920
Db	1861	CCTGCCACATATAGATGAGCCCATAGTTGGCCCCCTAGCTTCCAGCAGTCCACTATC	1920
QY	1921	TGCCAGAGGACCAAGGTGCCCTTAGACCTAAGCCAGGGGAAAGAGCATCTTCATAAAAA	1980
Db	1921	TGCCAGAGGACCAAGGTGCCCTTAGACCTAAGCCAGGGGAAAGAGCATCTTCATAAAAA	1980
QY	1981	CTTTCAGATCCAAACATTAATTTCTTTTATTTTCTGAGAGTTCAGGCAATCAGT	2040
Db	1981	CTTTCAGATCCAAACATTAATTTCTTTTATTTATCTGAGAGTTCAGGCAATCAGT	2040
QY	2041	ATTTCCCAAGGATGGGCAAGGCGCAGCCAGCGGCTTAGGATATCCAGCCTACCAAT	2100
Db	2041	ATTTCCCAAGGATGGGCAAGGCGCAGCCAGCGGCTTAGGATATCCAGCCTACCAAT	2100
QY	2101	ATGCTCATTCGACTAATCTAGAGGGTGTGTTGGCCCTGTCTCTTTTCTTGGACCTC	2160
Db	2101	ATGCTCATTCGACTAATCTAGAGGGTGTGTTGGCCCTGTCTCTTTTCTTGGACCTC	2160
QY	2161	AGTTTCTCTCAGTGAGTGGTAAAGATGCATTAACCTTTTGTATTTGATTAAGTTATAAATC	2220
Db	2161	AGTTTCTCTCAGTGAGTGGTAAAGATGCATTAACCTTTTGTATTTGATTAAGTTATAAATC	2220
QY	2221	TGTGGTCTGTATCTGAGTCCAGAGGGGAGATAGGTTCTGTGATTTTCTTCTTCTCT	2280
Db	2221	TGTGGTCTGTATCTGAGTCCAGAGGGGAGATAGGTTCTGTGATTTTCTTCTTCTCT	2280
QY	2281	ATAGAATAAATCAAACTTTGTTACTAGAAACAGAAATGTAGATGGCCCAAAACAAAGATG	2340
Db	2281	ATAGAATAAATCAAACTTTGTTACTAGAAACAGAAATGTAGATGGCCCAAAACAAAGATG	2340
QY	2341	ACCAGATTTGATCTCAGGCTGATGACCTCCTACAGGTCGTGTATGATGGAGTCTCATG	2400
Db	2341	ACCAGATTTGATCTCAGGCTGATGACCTCCTACAGGTCGTGTATGATGGAGTCTCATG	2400
QY	2401	GGTAAAGCAGGAAGAGAGTGGGAAAGAGAACCAACCCCACTCTCTTCATATTTTCATTT	2460
Db	2401	GGTAAAGCAGGAAGAGAGTGGGAAAGAGAACCAACCCCACTCTCTTCATATTTTCATTT	2460
QY	2461	CATGTTTAACTCCGCTGGAAATAGAAAGCAATCCCTTAGAGATGAGGATAAAGAAAG	2520
Db	2461	CATGTTTAACTCCGCTGGAAATAGAAAGCAATCCCTTAGAGATGAGGATAAAGAAAG	2520
QY	2521	TTTTCAGATTCACAGGGGAGAAATGAGAGATTTAATCCCTAAACCTGTGACTTGGGAG	2580

Db 2521 TTTCAGATTCACAGGGGGAAGAAATGAGATTAACTCTAAACCTGTGACTGGGAG 2580  
QY 2581 GTCAGCATTTTACAGTAGTCTGTGCTTTGACCTCTGTGATTATTAACCCACTCAC 2640  
Db 2581 GTCAGCATTTTACAGTAGTCTGTGCTTTGACCTCTGTGATTATTAACCCACTCAC 2640  
QY 2641 TACCTGTTCAGATGATTTGGAAATCCAAAGATTAATCCCTGATAGATTCAT 2700  
Db 2641 TACCTGTTCAGATGATTTGGAAATCCAAAGATTAATCCCTGATAGATTCAT 2700  
QY 2701 TGCAGAAAGCAGATTAAGAACATCAGAGAAATTAATTAGTGTATGACAGGCAA 2760  
Db 2701 TGCAGAAAGCAGATTAAGAACATCAGAGAAATTAATTAGTGTATGACAGGCAA 2760  
QY 2761 CTGTGAGAAACTGTGGCCAAATAATGAAATTCCTGTAGTTTCTGTGTCATTTGA 2820  
Db 2761 CTGTGAGAAACTGTGGCCAAATAATGAAATTCCTGTAGTTTCTGTGTCATTTGA 2820  
QY 2821 AAGAGAAATTCACATTTGTTAGCATTTCAAGCTTTTATGATCCATCCATCTAAA 2880  
Db 2821 AAGAGAAATTCACATTTGTTAGCATTTCAAGCTTTTATGATCCATCCATCTAAA 2880  
QY 2881 ACTCTTCAACTCCACTTGTTCAGTCTGAAATGCAAGTCTCCTGTCAGAGTCCCTGGAGA 2940  
Db 2881 ACTCTTCAACTCCACTTGTTCAGTCTGAAATGCAAGTCTCCTGTCAGAGTCCCTGGAGA 2940  
QY 2941 ACTCAGCAGCAGCCTTATCAAAAGGTTTACAGCCCTTGACACTATGAGAGAGG 3000  
Db 2941 ACTCAGCAGCAGCCTTATCAAAAGGTTTACAGCCCTTGACACTATGAGAGAGG 3000  
QY 3001 GCAAGAGTACACCAATTTGTTAAAAGCAAGAACCAAGTCTCTTCACTAGCATTTA 3060  
Db 3001 GCAAGAGTACACCAATTTGTTAAAAGCAAGAACCAAGTCTCTTCACTAGCATTTA 3060  
QY 3061 GAACATGTTATCATCAAGACTACTCTACCTGCAATTTGAATCCCAAGCAATC 3120  
Db 3061 GAACATGTTATCATCAAGACTACTCTACCTGCAATTTGAATCCCAAGCAATC 3120  
QY 3121 CACATTCCTGTGAGTTCTGAGCTTCTGTGTAATAGGAGCAGTGTGCTTATGCCCTA 3180  
Db 3121 CACATTCCTGTGAGTTCTGAGCTTCTGTGTAATAGGAGCAGTGTGCTTATGCCCTA 3180  
QY 3181 GAATACATGATCTGAGAGCAATTCATGGAAGCTGCTAATAGCTAGTCTGGGAGTCT 3240  
Db 3181 GAATACATGATCTGAGAGCAATTCATGGAAGCTGCTAATAGCTAGTCTGGGAGTCT 3240  
QY 3241 TCCATTAAGTTTGCATGAGAGCAAAACAGATTAACCTAGTTGGTCTTCAAGCC 3300  
Db 3241 TCCATTAAGTTTGCATGAGAGCAAAACAGATTAACCTAGTTGGTCTTCAAGCC 3300  
QY 3301 CTCTTAAGCATAGGCTTATGCTGAGAGCTTCTGTGGGCTTCTGTGTTAGTTT 3360  
Db 3301 CTCTTAAGCATAGGCTTATGCTGAGAGCTTCTGTGGGCTTCTGTGTTAGTTT 3360  
QY 3361 TGTAAACACTATAGATCTGTTAAGATCCAGTGCATGAAACCTTCCCATAGCGTG 3420  
Db 3361 TGTAAACACTATAGATCTGTTAAGATCCAGTGCATGAAACCTTCCCATAGCGTG 3420  
QY 3421 ACTCTGAGCTATATCAAGTTTGGAAAGAGGCTTCTGTGCTGCTGCTTAAACAGCCAGT 3480  
Db 3421 ACTCTGAGCTATATCAAGTTTGGAAAGAGGCTTCTGTGCTGCTGCTTAAACAGCCAGT 3480  
QY 3481 GGAACAGTCTGATGCTTCTTCTTACACTATGTTTAAAGTACTCAAACTTCAAGAA 3540  
Db 3481 GGAACAGTCTGATGCTTCTTCTTACACTATGTTTAAAGTACTCAAACTTCAAGAA 3540  
QY 3541 CAATCTAAACAAGTTTCTGTGATATGTTTGGAACTTGTATTTATTTAGTGGC 3600  
Db 3541 CAATCTAAACAAGTTTCTGTGATATGTTTGGAACTTGTATTTATTTAGTGGC 3600  
QY 3601 TTTCTATATGCAATTTACTGTTTGTGTAACCTGATTTCTTCTTGGATCTATG 3660  
Db 3601 TTTCTATATGCAATTTACTGTTTGTGTAACCTGATTTCTTCTTGGATCTATG 3660

QY 3661 ATGAATTAAGAAATTAAGTGAAGAAAAA 3693  
Db 3661 ATGAATTAAGAAATTAAGTGAAGAAAAA 3693

RESULT 3  
US-09-880-107-3287  
; Sequence 3287; Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Home, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-MO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 3287  
; LENGTH: 1987  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U06863  
US-09-880-107-3287

Query Match 52.5%; Score 1940; DB 9; Length 1987;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1954; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 12 CTCGCCCTCGGCTTACAGCTGCGCGCGCTCGTCCGCCGCGCGCCAGAGACTGG 71  
Db 21 CTCGCCCTCGGCTTACAGCTGCGCGCGCTCGTCCGCCGCGCGCCAGAGACTGG 80  
QY 72 ACCAGACAGATGTGAAACGCTGCTGCGCTGCGCTGCGCTGCTGCTGCTGCT 131  
Db 81 ACCAGACAGATGTGAAACGCTGCTGCGCTGCGCTGCGCTGCTGCTGCTGCT 140  
QY 132 GGGTCCGCCCGAGAGAGCTAAGAGCAATCCAAAGTCTGTCGAATGTTTG 191  
Db 141 GGGTCCGCCCGAGAGAGCTAAGAGCAATCCAAAGTCTGTCGAATGTTTG 200  
QY 192 GAGCCGCGCGGAGATGTGCACTCAGAGAAAGGAAACCACTGTCTCTGATGAGC 251  
Db 201 GAGCCGCGCGGAGATGTGCACTCAGAGAAAGGAAACCACTGTCTCTGATGAGC 260  
QY 252 AATGCAAACTCAGAAAGAGCTGTGTGAGCAATAGGCAAGCACTCAACCACT 311  
Db 261 AATGCAAACTCAGAAAGAGCTGTGTGAGCAATAGGCAAGCACTCAACCACT 320  
QY 312 GTGAATGATGAGATGCTGCTCACTGATCCAAATCCAGGTTGATTCATGAGAC 371  
Db 321 GTGAATGATGAGATGCTGCTCACTGATCCAAATCCAGGTTGATTCATGAGAC 380  
QY 372 ACTGCAAGAGAAATTCGTAAGTCCATCTGCCAGGCCAGTTGTTGCTATCAGTCCA 431  
Db 381 ACTGCAAGAGAAATTCGTAAGTCCATCTGCCAGGCCAGTTGTTGCTATCAGTCCA 440  
QY 432 ACCGTATGAGCTCCAGCTGCGATCATCAAGTGTGGAAGCTGATCATCTAGATG 491  
Db 441 ACCGTATGAGCTCCAGCTGCGATCATCAAGTGTGGAAGCTGATCATCTAGATG 500  
QY 492 GCTGCTTCTTAAAGCAGCACTAGTGAATCTTGAACAATTTTAAAGACTTGG 551  
Db 501 GCTGCTTCTTAAAGCAGCACTAGTGAATCTTGAACAATTTTAAAGACTTGG 560

552 ATAAATGGTGAATTCCTGGCTGCACTCCAGTGAATTCCTGAAGTTTGTGGACAGATGAAA 611  
561 ATAAATGGTGAATTCCTGGCTGCACTCCAGTGAATTCCTGAAGTTTGTGGACAGATGAAA 620  
612 CTGCCATCATATATACAACTGATCCAGACAGAGAACAACTGCTTAGGGGACTCT 671  
621 CTGCCATCATATATACAACTGATCCAGACAGAGAACAACTGCTTAGGGGACTCT 680  
672 GTGTGATGCTCTCATTTGAATCTGCTGATGAAAATGCTGATTCGAAATCAGCTTCCAAG 731  
681 GTGTGATGCTCTCATTTGAATCTGCTGATGAAAATGCTGATTCGAAATCAGCTTCCAAG 740  
732 AGTTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCTGAGAGAAAGTGTGCCCTGGAG 791  
741 AGTTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCTGAGAGAAAGTGTGCCCTGGAG 800  
792 ATGAAACGATGATGAGATGGAGCTGAGACCGAGGTGGACTCTAAACCGCTGTGTGTGCCT 851  
801 ATGAAACGATGATGAGATGGAGCTGAGACCGAGGTGGACTCTAAACCGCTGTGTGTGCCT 860  
852 GTGAAATTTGGTCTGTACAGCCATGACCTGTGACCGAAGAGATCAGAGGGGGCCAG 911  
861 GTGAAATTTGGTCTGTACAGCCATGACCTGTGACCGAAGAGATCAGAGGGGGCCAG 920  
912 CCCAGACAGAGGAGGATGACCAAGATATGTCCAGGAGCTCCAAAAGCATCAGGAAACAG 971  
921 CCCAGACAGAGGAGGATGACCAAGATATGTCCAGGAGCTCCAAAAGCATCAGGAAACAG 980  
972 CTGAAAAGACCAAGAGTGAAGACCAAGAGATCTAATGAGGAGCACAGACCATGTC 1031  
981 CTGAAAAGACCAAGAGTGAAGACCAAGAGATCTAATGAGGAGCACAGACCATGTC 1040  
1032 TGGATCCAGCATCTTCTCCACTTCAGCGTGTGAGTTCAGTATACAAAGTGTCTGTACA 1091  
1041 TGGATCCAGCATCTTCTCCACTTCAGCGTGTGAGTTCAGTATACAAAGTGTCTGTACA 1100  
1092 GTCCGCAATCAGCATTTTGTCTTATAGCAATGATTTTATTTTATTTTATTTT 1151  
1101 GTCCGCAATCAGCATTTTGTCTTATAGCAATGATTTTATTTTATTTTATTTT 1160  
1152 GCAATAAGAGTATGAAGTGGCTGGCTAGGAAGGAGGCGCACAGCTTCAATTTCTAG 1211  
1161 GCAATAAGAGTATGAAGTGGCTGGCTAGGAAGGAGGCGCACAGCTTCAATTTCTAG 1220  
1212 GAGTGTCTTAAAGAGAACTGTAAATGTGTCTCTGGGCTGGAGCTAGTAAGAACTGC 1271  
1221 GAGTGTCTTAAAGAGAACTGTAAATGTGTCTCTGGGCTGGAGCTAGTAAGAACTGC 1280  
1272 ATCAGGATTGAAGAGGAGAACCAATCTGAACCTCTTTGAGTTTACTGCTATCTGT 1331  
1281 ATCAGGATTGAAGAGGAGAACCAATCTGAACCTCTTTGAGTTTACTGCTATCTGT 1340  
1332 CAGCAGGCTGAGGAGTGCACAGATGCCAGAGAGAACTTAGCAGGTTTCCCGGAGG 1391  
1341 CAGCAGGCTGAGGAGTGCACAGATGCCAGAGAGAACTTAGCAGGTTTCCCGGAGG 1400  
1392 AGAGTTTGGGAGCTCCACGAGGAGGAGCTCTCTGCTCCAGCTCTTTCCATTGCC 1451  
1401 AGAGTTTGGGAGCTCCACGAGGAGGAGCTCTCTGCTCCAGCTCTTTCCATTGCC 1460  
1452 GTCAGCATGACAGACCTCCAGCATCCAGCATCTCTTGTGCCAATTAATGCCTCTAGAT 1511  
1461 GTCAGCATGACAGACCTCCAGCATCCAGCATCTCTTGTGCCAATTAATGCCTCTAGAT 1520  
1512 ACATAGCATACTGTAGTTAACCNGTGTCCCTCAGACTTTGGATGGATTTCTGGGAGG 1571  
1521 ACATAGCATACTGTAGTTAACCNGTGTCCCTCAGACTTTGGATGGATTTCTGGGAGG 1580  
1572 GTACACCCAAATGATGACAGATCTTGTATCTTTGAGCCCTTAGCGACCTTAACCAATT 1631  
1581 GTACACCCAAATGATGACAGATCTTGTATCTTTGAGCCCTTAGCGACCTTAACCAATT 1640  
1632 TTAATAATACCTTTTACCAAGGCTGCTATTTCTCTGTAAACACATTTTTTTTGGCAAGTT 1691

Db 1641 TTAATAATACCTTTTACCAAGGCTGCTATTTCTCTGTAAACAC-TTTTTTGGCAAGTT 1699  
Qy 1692 GACTTTATCTTCAATTAATATCAATTAATATTTTAAATTAATTTTCTTGAC 1751  
Db 1700 GACTTTATCTTCAATTAATATCAATTAATATTTTAAATTAATTTTCTTGAC 1759  
Qy 1752 TAGTATTAAGCTTTTCTGTAATTAATTTTCACTAGTCCACCACTTCATAGTGGAGGAG 1811  
Db 1760 TAGTATTAAGCTTTTCTGTAATTAATTTTCACTAGTCCACCACTTCATAGTGGAGGAG 1819  
Qy 1812 TTTGGGTTCTTCTGTTGTCAGGGGCTGAAATAACCCAGATGCTCCACCTGCCACATA 1871  
Db 1820 TTTGGGTTCTTCTGTTGTCAGGGGCTGAAATAACCCAGATGCTCCACCTGCCACATA 1879  
Qy 1872 CTAGATGAGCCCATAGTTGGCCCCCTAGCTTCCAGACTCCACTATCTGCCAGAGGAG 1931  
Db 1880 CTAGATGAGCCCATAGTTGGCCCCCTAGCTTCCAGACTCCACTATCTGCCAGAGGAG 1939  
Qy 1932 CAAGGTCCTTTAGACCTTAAGCCAGGGGAGGAAGCATCTT 1971  
Db 1940 CAAGGTCCTTTAGACCTTAAGCCAGGGGAGGAAGCATCTT 1979

## RESULT 4

US-09-960-706-941  
; Sequence 941, Application US/09960706  
; Publication No. US20030134280A1  
; GENERAL INFORMATION:  
; APPLICANT: Munger, William E.  
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia  
; FILE REFERENCE: 44921-5029-0105  
; CURRENT APPLICATION NUMBER: US/09/960,706  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 60/223,323  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: 09/873,319  
; PRIOR FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 1124  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 941  
; LENGTH: 1987  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 U06863  
US-09-960-706-941

Query Match 52.5%; Score 1940; DB 10; Length 1987;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1954; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
Qy 12 CTCCTCACTTCGCTTACAGCTCGCTGCCCGCTGCTGCCCGCGCCGCCCGCCCGCCAGAGACCTGG 71  
Db 21 CTCCTCACTTCGCTTACAGCTCGCTGCCCGCTGCTGCCCGCGCCGCCCGCCCGCCAGAGACCTGG 80  
Qy 72 ACCAGACCAACGATGTGGAACCGCTGCGCTCGCGCTCGCGCTCGCGCTGCGCTGCGCTGCGCT 131  
Db 81 ACCAGACCAACGATGTGGAACCGCTGCGCTCGCGCTCGCGCTCGCGCTGCGCTGCGCTGCGCT 140  
Qy 132 GGGTCGGGCGGAGGAGGAGCTAAGAGCAATCCAAATCTGTGCCAATGTGTTTGTG 191  
Db 141 GGGTCGGGCGGAGGAGGAGCTAAGAGCAATCCAAATCTGTGCCAATGTGTTTGTG 200  
Qy 192 GAGCGGCGGCGGAAATGTGAGTCAAGAGAAAGGGGAAACCCACCTGTCTCTGCAATTGAGC 251  
Db 201 GAGCGGCGGCGGAAATGTGAGTCAAGAGAAAGGGGAAACCCACCTGTCTCTGCAATTGAGC 260  
Qy 252 AATGCAAACTCACAAGAGGCTGTGTGTGGAGTAATGGCAAGACCTTACCTCAACCACT 311  
Db 261 AATGCAAACTCACAAGAGGCTGTGTGTGGAGTAATGGCAAGACCTTACCTCAACCACT 320

QY	312	GTGAACTGCATCGAGATGCTGGCTCAGTGGATCCAAATCCAGGTGATTAAGATGAGAC	371
Db	321	GTGAACTGCATCGAGATGCTGGCTCAGTGGATCCAAATCCAGGTGATTAAGATGAGAC	380
QY	372	ACTGCAAGAGAAATCCGTAACTGCTCCAGCCGATTTTGAACCTTG	431
Db	381	ACTGCAAGAGAAATCCGTAACTGCTCCAGCCGATTTTGAACCTTG	440
QY	432	ACCGTATGAGCTCCGACGTGGATCATCCAGTGGCTGAGAGGCTGAGATCATCTCGATG	491
Db	441	ACCGTATGAGCTCCGACGTGGATCATCCAGTGGCTGAGAGGCTGAGATCATCTCGATG	500
QY	492	GCTGTTCTCTAAAGCAGCACTACAGTAAATCTTACAAATTTTGAACCTTG	551
Db	501	GCTGTTCTCTAAAGCAGCACTACAGTAAATCTTACAAATTTTGAACCTTG	560
QY	552	ATATATGATATTTCTGGCTGATCTCCAGTAAATTTCTGAAAGTTTGGGAAACAGATGAA	611
Db	561	ATATATGATATTTCTGGCTGATCTCCAGTAAATTTCTGAAAGTTTGGGAAACAGATGAA	620
QY	612	CTGCAATCATATTATCAACGTATCCAGACCGAGAGAAACAAGTTGCTTAAAGGACCTCT	671
Db	621	CTGCAATCATATTATCAACGTATCCAGACCGAGAGAAACAAGTTGCTTAAAGGACCTCT	680
QY	672	GTGTTGATGCTCTTCATTGAACTGTCTGATGAAAACTGATTTGAAACTCAGCTTCCAG	731
Db	681	GTGTTGATGCTCTTCATTGAACTGTCTGATGAAAACTGATTTGAAACTCAGCTTCCAG	740
QY	732	AGTTTCTCAAGTGGCTCAACCCATCTTCAACCTCCTGAGAAAGTGTGACCTTGAGG	791
Db	741	AGTTTCTCAAGTGGCTCAACCCATCTTCAACCTCCTGAGAAAGTGTGACCTTGAGG	800
QY	792	ATGAAACGTATGCAATGAGACTGAGACCGAGGTGACCTGTAAACCGCTGTCTGTGCT	851
Db	801	ATGAAACGTATGCAATGAGACTGAGACCGAGGTGACCTGTAAACCGCTGTGTCTGCT	860
QY	852	GTGGAATTTGGGTCTGTACAGCCATGTGACCGGAAAGAAATCAAGAGGGGCCAG	911
Db	861	GTGGAATTTGGGTCTGTACAGCCATGTGACCGGAAAGAAATCAAGAGGGGCCAG	920
QY	912	CCCAGACAGAGAGAGATGACACGATATGTCCAGAGCTCCAAAGCATCAGAAACAG	971
Db	921	CCCAGACAGAGAGAGATGACACGATATGTCCAGAGCTCCAAAGCATCAGAAACAG	980
QY	972	CTGAAAGAACCAAGAGTGAACCCAAAGATCTTAATGAGAGGCACACAGCAAGTCT	1031
Db	981	CTGAAAGAACCAAGAGTGAACCCAAAGATCTTAATGAGAGGCACACAGCAAGTCT	1040
QY	1032	TGATATCCAGATCTTCTCACTCAGGCTGAGATTCAATTAACAAGTGTCTGTACA	1091
Db	1041	TGATATCCAGATCTTCTCACTCAGGCTGAGATTCAATTAACAAGTGTCTGTACA	1100
QY	1092	GTGCGCAATCACAAGTATTTGCTTAATACCAATGATTTTATTTTATTTGTTTT	1151
Db	1101	GTGCGCAATCACAAGTATTTGCTTAATACCAATGATTTTATTTTATTTGTTTT	1160
QY	1152	GCAATTAAGATTAAGAGTGGCTGGCTAGGAAAGGACCAAGCTCATTTCTAG	1211
Db	1161	GCAATTAAGATTAAGAGTGGCTGGCTAGGAAAGGACCAAGCTCATTTCTAG	1220
QY	1212	GAGTGTCTTTAAGAGAACTGTAAATGTGTCTCTGGGAGTGAAGCTAGTAGAGAACTGC	1271
Db	1221	GAGTGTCTTTAAGAGAACTGTAAATGTGTCTCTGGGAGTGAAGCTAGTAGAGAACTGC	1280
QY	1272	ATCAGCATTTGAAAGAGAAACAGACCCAAATCTTAACTTTTGAATTTACATCTGT	1331
Db	1281	ATCAGCATTTGAAAGAGAAACAGACCCAAATCTTAACTTTTGAATTTACATCTGT	1340
QY	1332	CAGCAGGCTGACAGGAGTGCACACGATCCAGAGAGAACTTAGCAGGGTGTCCCGAGAG	1391
Db	1341	CAGCAGGCTGACAGGAGTGCACACGATCCAGAGAGAACTTAGCAGGGTGTCCCGAGAG	1400
QY	1392	AGAGGTTTGGGAAGCTCCACGGAGAGAAAGCTCTGTGCTTCCAGCTTTTCCATTGGC	1451

Db	1401	AGAGGTTTGGAGAGCTCCAGAGAGAGAGAGCGCTCTGCTTCCAGCCTCTTCCATGGCC	1460
Qy	1452	GTGAGCATGACAGACCTCCAGCATCCAGCATCTCTTGCTCCAAATACAGCTGCTTAGAT	1511
Db	1461	GTGAGCATGACAGACCTCCAGCATCCAGCATCTCTTGCTCCAAATACAGCTGCTTAGAT	1520
Qy	1512	ACATAGCACAACGCTAGTAAACCAACGAGTGCCTCCAGACTTGGATGGAGTTTCTGGAGG	1571
Db	1521	ACATAGCACAACGCTAGTAAACCAACGAGTGCCTCCAGACTTGGATGGAGTTTCTGGAGG	1580
Qy	1572	GTACACCCAAATGATGAGATGACTTGTATACTTTGAGCCCTTAGCGACTAACCAATT	1631
Db	1581	GTACACCCAAATGATGAGATGACTTGTATACTTTGAGCCCTTAGCGACTAACCAATT	1640
Qy	1632	TTTAAATACTCTTTTACCAAGAGGATTTCTGTGAAACACTTTTCTTTGGCAAGT	1691
Db	1641	TTTAAATACTCTTTTACCAAGAGGATTTCTGTGAAACAC-TTTTCTTTGGCAAGT	1699
Qy	1692	GACTTATCTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTCTTGAC	1751
Db	1700	GACTTATCTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTCTTGAC	1759
Qy	1752	TAGTATTAAGCTTTTGTAATTAATTTTCAAGTATCCACACACTCAATAGTGGAGAGG	1811
Db	1760	TAGTATTAAGCTTTTGTAATTAATTTTCAAGTATCCACACACTCAATAGTGGAGAGG	1819
Qy	1812	TTTGGAGGTTCTCTCGTGTGAGGGGCTGAATATACCAGATGCTCCACACTGCGACATA	1871
Db	1820	TTTGGAGGTTCTCTCGTGTGAGGGGCTGAATATACCAGATGCTCCACACTGCGACATA	1879
Qy	1872	CTAGATCAGAGCCCATTAATTGAGCCCTTACGTTCCAGAGTCCACTATCTGCGAGAGAG	1931
Db	1880	CTAGATCAGAGCCCATTAATTGAGCCCTTACGTTCCAGAGTCCACTATCTGCGAGAGAG	1939
Qy	1932	CAAGGAGGCTTAGACTTAAGCCAGGAGGAGAAAGACATCTT	1971
Db	1940	CAAGGAGGCTTAGACTTAAGCCAGGAGGAGAAAGACATCTT	1979

```

RESULT 5
US-09-873-319-617
: Sequence 617, Application US/09873319A
: Publication No. US20030134324A1
GENERAL INFORMATION:
APPLICANT: Munger, William E.
APPLICANT: Kulkarni, Prakash
APPLICANT: Getzenberg, Robert H.
APPLICANT: Waga, Iwao
APPLICANT: Yamamoto, Jun
: TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
: TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
: FILE REFERENCE: 44921-5029-US
CURRENT APPLICATION NUMBER: US/09/873.319A
CURRENT FILING DATE: 2001-06-05
EARLIER APPLICATION NUMBER: US 60/223.323
EARLIER FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 755
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 617
: LENGTH: 1987
: TYPE: DNA
: ORGANISM: Homo sapiens
FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20030134324A1 U06863
US-09-873-319-617

Query Match          52.5%  Score 1940;  DB 10;  Length 1987;
Best Local Similarity 99.7%  Pred. No. 0;
Matches 1954;  Conservative 0;  Mismatches 5;  Indels 1;  Gaps 1.
12 CTCGCACCTCGGCTTACAGCTCGCTGCGCGCGCTCTGCGCGCGCGCCAGAGACCTGG 71

```

Db 21 CTCACACTCGCTTACAGCTCGCTGCGCGCTCTCTGCTCCCGCGCCCGCCAGGAGACTCG 80  
Qy 72 ACAGACACAGATGTGAACGCTGGCTCGGCTCGGCTCGGCTCGGCTGGTGGCGGTGCGCT 131  
Db 81 ACCAGACACAGATGTGAACGCTGGCTCGGCTCGGCTCGGCTCGGCTGGTGGCGGTGCGCT 140  
Qy 132 GGGTCCGCGCGCGAGAGAGCTAAGGAGCAAAATCCAGATCTGTGCCAAATGTGTTTGTG 191  
Db 141 GGGTCCGCGCGCGAGAGAGCTAAGGAGCAAAATCCAGATCTGTGCCAAATGTGTTTGTG 200  
Qy 192 GAGCGCGCGGGAATGTGAGTCAAGAGAAAGGGAACCCACTGTGTCTGTGCAATGAGC 251  
Db 201 GAGCGCGCGGGAATGTGAGTCAAGAGAAAGGGAACCCACTGTGTCTGTGCAATGAGC 260  
Qy 252 AATGCAAACTCACAAGAGGCTGTGTGTGCGCAGTAAATGCAAGACTACTCAACCACT 311  
Db 261 AATGCAAACTCACAAGAGGCTGTGTGTGCGCAGTAAATGCAAGACTACTCAACCACT 320  
Qy 312 GTGAATGCAATCGAGATGCTGTGCTCTACTGATCCAAATCCAGGTTGATTAAGTGGAC 371  
Db 321 GTGAATGCAATCGAGATGCTGTGCTCTACTGATCCAAATCCAGGTTGATTAAGTGGAC 380  
Qy 372 ACTGCAAGAGAGAAATCGTAAAGTCCATCTGCGAGCCAGTGTGTTGCTATCAGTCCA 431  
Db 381 ACTGCAAGAGAGAAATCGTAAAGTCCATCTGCGAGCCAGTGTGTTGCTATCAGTCCA 440  
Qy 432 ACCGTGATGACTCCGAGTCCGATCATCCAGTGGCTGGAAGCTGAGATCATCTCAGATG 491  
Db 441 ACCGTGATGACTCCGAGTCCGATCATCCAGTGGCTGGAAGCTGAGATCATCTCAGATG 500  
Qy 492 GCTGGTCTCTAAGAGGAGCAACTACAGTGAATCTTAGCAAGTATTTTAAAGACTTTG 551  
Db 501 GCTGGTCTCTAAGAGGAGCAACTACAGTGAATCTTAGCAAGTATTTTAAAGACTTTG 560  
Qy 552 ATAAATGAGTCTCGCTGAGTCCAGTGAATCTTGAAGTGTGGAACAGAAATGAAA 611  
Db 561 ATAAATGAGTCTCGCTGAGTCCAGTGAATCTTGAAGTGTGGAACAGAAATGAAA 620  
Qy 612 CTGCGATCAATATACAGTATCCAGACAGGAGACACAGTGTCTTAGGGGACTCT 671  
Db 621 CTGCGATCAATATACAGTATCCAGACAGGAGACACAGTGTCTTAGGGGACTCT 680  
Qy 672 GTGTGATGCTCTCATTAAGTCTGATGAAATGCTGATGGAACACTCAGTCTCCAAG 731  
Db 681 GTGTGATGCTCTCATTAAGTCTGATGAAATGCTGATGGAACACTCAGTCTCCAAG 740  
Qy 732 AGTTCTCAAGTGCCTCAACCCATCTTCAACCCCTCTGAGAGAAAGTGTGCTGGAGG 791  
Db 741 AGTTCTCAAGTGCCTCAACCCATCTTCAACCCCTCTGAGAGAAAGTGTGCTGGAGG 800  
Qy 792 ATGAAAAGTATGAGATGGAGCTGAGACCGAGTGGACTGTAAACCGCTGTGTCTGTGCT 851  
Db 801 ATGAAAAGTATGAGATGGAGCTGAGACCGAGTGGACTGTAAACCGCTGTGTCTGTGCT 860  
Qy 852 GTGGAATTTGGTCTGTACAGCCATGACCTGTGACGGAAGATCAGAAAGGCGCCAGAG 911  
Db 851 GTGGAATTTGGTCTGTACAGCCATGACCTGTGACGGAAGATCAGAAAGGCGCCAGAG 920  
Qy 912 CCCAGACAGAGAGAGATGACAGATATGTCAGGAGCTCCAAAGCATCAGGAAACAG 971  
Db 921 CCCAGACAGAGAGAGATGACAGATATGTCAGGAGCTCCAAAGCATCAGGAAACAG 980  
Qy 972 CTGAAAAGACCAAGAGATGAGCCAAAGAGATCTAATGAGGAGGACACAGACAGTGTG 1031  
Db 981 CTGAAAAGACCAAGAGATGAGCCAAAGAGATCTAATGAGGAGGACACAGACAGTGTG 1040  
Qy 1032 TGGATCCAGCATCTTCTCCACTTCCAGGCTGAGTTCAGTATACACAAAGTGTCTGTACA 1091  
Db 1041 TGGATCCAGCATCTTCTCCACTTCCAGGCTGAGTTCAGTATACACAAAGTGTCTGTACA 1100  
Qy 1092 CTCGCAAAATCACCAGTATTTGCTTATAGCAATGAGTTTATTTTGTATTTTGT 1151  
Db 1101 CTCGCAAAATCACCAGTATTTGCTTATAGCAATGAGTTTATTTTGTATTTTGT 1160

Qy 1152 GCATTAAGGATATGATGAGTGGCTGGCTAGGAAGGAGGCGCACAGCCTTCATTTCTAG 1211  
Db 1161 GCATTAAGGATATGATGAGTGGCTGGCTAGGAAGGAGGCGCACAGCCTTCATTTCTAG 1220  
Qy 1212 GAGTCTTTTAAGAGAAACTGTAAATGTGTCTCTGGGCTGGAGCTAGTAAGGAAACTGC 1271  
Db 1221 GAGTCTTTTAAGAGAAACTGTAAATGTGTCTCTGGGCTGGAGCTAGTAAGGAAACTGC 1280  
Qy 1272 ATCAGATTTGAAGAGGAAACAGACCCAAATCTGAACCTCTTTTCTGAGTTTACTGATCTGT 1331  
Db 1281 ATCAGATTTGAAGAGGAAACAGACCCAAATCTGAACCTCTTTTCTGAGTTTACTGATCTGT 1340  
Qy 1332 CAGCAGGCTGAGGAGTGCACAGATGCCAGAGAGAACTTTAGCAGGCTGTCCCGGAGG 1391  
Db 1341 CAGCAGGCTGAGGAGTGCACAGATGCCAGAGAGAACTTTAGCAGGCTGTCCCGGAGG 1400  
Qy 1392 AGAGTTTGGGAAAGCTCCACGAGAGGAAACCTCTCTGCTTCCAGCCTCTTTTCCATTTGCC 1451  
Db 1401 AGAGTTTGGGAAAGCTCCACGAGAGGAAACCTCTCTGCTTCCAGCCTCTTTTCCATTTGCC 1460  
Qy 1452 GTCAGATGACAGACCTCCAGCATCCAGCATCCAGCATCTCTTGGTCCCAATTAACCTCTTAGAT 1511  
Db 1461 GTCAGATGACAGACCTCCAGCATCCAGCATCTCTTGGTCCCAATTAACCTCTTAGAT 1520  
Qy 1512 ACATAGCATATCTGTAGTTAAACCAAGTGTCCCTCAGACTTGGATGGAGTTTCTGGGAGG 1571  
Db 1521 ACATAGCATATCTGTAGTTAAACCAAGTGTCCCTCAGACTTGGATGGAGTTTCTGGGAGG 1580  
Qy 1572 GTACACCCAAATGATGACATCTGTATATCTTTGAGCCCTTAGCGACCTAAACCAAT 1631  
Db 1581 GTACACCCAAATGATGACATCTGTATATCTTTGAGCCCTTAGCGACCTAAACCAAT 1640  
Qy 1632 TTAAAAATCTTTTACCAAAAGTGTATTTCTCTGTAAAAACACTTTTTTTTGGCAAGTT 1691  
Db 1641 TTAAAAATCTTTTACCAAAAGTGTATTTCTCTGTAAAAACACTTTTTTTTGGCAAGTT 1699  
Qy 1692 GACTTTATCTTCAATTTCTTGAC 1751  
Db 1700 GACTTTATCTTCAATTTCTTGAC 1759  
Qy 1752 TAGTATTAAGCTTTTGTAAATATTTTTCAGTATGCCACACTTCATAGTGGAGGAG 1811  
Db 1760 TAGTATTAAGCTTTTGTAAATATTTTTCAGTATGCCACACTTCATAGTGGAGGAG 1819  
Qy 1812 TTTGGGTTCTTCTGTCGAGGCTGAAATAACCCAGATGCTCCACCTGCGCACATA 1871  
Db 1820 TTTGGGTTCTTCTGTCGAGGCTGAAATAACCCAGATGCTCCACCTGCGCACATA 1879  
Qy 1872 CTAGATGAGCCCATAGTTGGCCCCCTTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG 1931  
Db 1880 CTAGATGAGCCCATAGTTGGCCCCCTTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG 1939  
Qy 1932 CAAGGTCCTTAGACCTAAGCCAGGGAAGCAAGCATCTT 1971  
Db 1940 CAAGGTCCTTAGACCTAAGCCAGGGAAGCAAGCATCTT 1979

RESULT 6  
US-10-098-841-144/c  
; Sequence 144, Application US/10098841  
; Publication No. US20020197679A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Xu, Chongjun  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yungqing  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong







Db	970	TCACCTATCTCCAAATTATATACATAATATATGTTTTTAATATTTATTTCTTGA	911
QY	1751	CTAGG	1755
Db	910	CTAGG	906

```

RESULT 7
US-09-764-877-3903
: Sequence 3903, Application US/09764877
: Patent No US20020147710A1
GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC005
: CURRENT APPLICATION NUMBER: US/09/764,877
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 4031
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3903
: LENGTH: 1755
: TYPE: DNA
: ORGANISM: Homo sapiens
FEATURE:
: NAME/KEY: SITE
: LOCATION: (2)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (4)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-764-877-3903

```

Query Match	45.4%	Score 1678.4	DB 9	Length 1755
Best Local Similarity	99.5%	Pred. No. 0		
Matches 1682; Conservative	0	Mismatches 8	Indels 0	Gaps 0

QY	1993	AAACCTAAATGGTTTATTTATTTATCTGAGAAGTGAAGCAAAATCAGATTTCCCAAGAT	2055
Db	1	ANANNTTTTGTGTTTTTTTATTTATTTCTGGAAGTTGAGGCANAATGATATTTCCAAAGAT	60
QY	2053	GGCGCAAGGCGACGCCAAGCAGGGCTTAGATATCCAGCCTCAAAATATGCTCATTTGCA	2112
Db	61	GGCGCAAGGCGACGCCAAGCAGGGCTTAGATATCCAGCCTCAAAATATGCTCATTTGCA	120
QY	2113	CTAACTAGAGGGTGAGTTGCGCCCGTCTCTTCTTTTCTGGAACCTAGTTCCCTGAT	2172
Db	121	CTAACTAGAGGGTGAGTTGCGCCCGTCTCTTCTTTTCTGGAACCTAGTTCCCTGAT	180
QY	2173	GAGCTGGTAGAATGCACATTAACCTTTGATTTGATTAAGTTTAAATTCGTGGTTCGAT	2232
Db	181	GAGCTGGTAGAATGCACATTAACCTTTGATTTGATTAAGTTTAAATTCGTGGTTCGAT	240
QY	2233	CATTGGTCCAGGGGAGATAGGTCCGCTGATTTTTCCTCTCTCTATAGAAATTAATG	2292
Db	241	CATTGGTCCAGGGGAGATAGGTCCGCTGATTTTTCCTCTCTCTATAGAAATTAATG	300
QY	2293	AAATCTTGTTACTAGAACCAAGAAATGTCCAGATGGCCAAAAAACAAGATGACCAAGATTGAT	2352
Db	301	AAATCTTGTTACTAGAACCAAGAAATGTCCAGATGGCCAAAAAACAAGATGACCAAGATTGAT	360
QY	2353	CTCAGCCCTGATGACCCCTACAGGTCCGTGCTATGATATGGAAGTCCCTATGGGTAAAGCAGA	2412
Db	361	CTCAGCCCTGATGACCCCTACAGGTCCGTGCTATGATATGGAAGTCCCTATGGGTAAAGCAGA	420
QY	2413	AGAGAAGTGGAAAAGAAACCAACCCCACTCTGTCTTCATATTTGCAATTCATGTTTAACT	2472
Db	421	AGAGAAGTGGAAAAGAAACCAACCCCACTCTGTCTTCATATTTGCAATTCATGTTTAACT	480
QY	2473	CCGGCTGGAAATPAGAAAGCATTTCCCTTGAGATAGAGTAAGAAAGAAATTCCTCAATTCAA	2532
Db	481	CCGGCTGGAAATPAGAAAGCATTTCCCTTGAGATAGAGTAAGAAAGAAATTCCTCAATTCAA	540

QY	2533	AAGGGGGGAAAGAAATGAGATTTAAATCTTAAACCTGGACTTGGGGAGGTCAGCATTTA	2532
Db	541	CAGGGGGGAAAGAAATGAGATTTAAATCTTAAACCTGGACTTGGGGAGGTCAGCATTTA	600
QY	2593	CAGTTAGTCTGTGTCCTTTGCACTTCTGTGATTTTAAACCCCACTCACTACCTCTTTGA	2652
Db	601	CAGTTAGTCTGTGTCCTTTGCACTTCTGTGATTTTAAACCCCACTCACTACCTCTTTGA	660
QY	2653	GATGCATTTGGAAATACCAAGAATTAATAATCCCTGACATAAAGTCATTTGGAGAAAGAG	2712
Db	661	GATGCATTTGGAAATACCAAGAATTTAAATCTTGACATAAAGTCATTTGGAGAAAGAG	720
QY	2713	ATTAAAGACATCAAGAAAGAAATTAATTTAGTTGTAAATGCAACAGCAACTGTGAAACT	2772
Db	721	ATTAAAGACATCAAGAAAGAAATTAATTTAGTTGTAAATGCAACAGCAACTGTGAAACT	780
QY	2773	GTTGNGCAAAAATAGAAATCCCTCTAGTTTCTTTGTTCTCATTTGAAAGAGAAATTT	2832
Db	781	GTTGNGCAAAAATAGAAATTCCTCTAGTTTCTTTGTTCTCATTTGAAAGAGAAATTT	840
QY	2833	CCACTTTGTTTAGCATTTCAAAGCTTTATGTATTCATCCCATCTAAAACTCTTCAAAC	2892
Db	841	CCACTTTGTTTAGCATTTCAAAGCTTTATGTATTCATCCCATCTAAAACTCTTCAAAC	900
QY	2893	CCAATTGTTCAGTCTGAAAATGCAAGCTGCTCCGTCCAAGTGCCTTGGAGAACTCACAGAC	2952
Db	901	CCAATTGTTCAGTCTGAAAATGCAAGCTGCTCCGTCCAAGTGCCTTGGAGAACTCACAGAC	960
QY	2953	ACGCTTAAATCAAAAGTTTTTACCAAGCCTTGGACACTATGGAGAGAGGCAAGATACAC	3012
Db	961	ACGCTTAAATCAAAAGTTTTTACCAAGCCTTGGACACTATGGAGAGAGGCAAGATACAC	1020
QY	3013	CAATTTGTTAAAAGCAAGAAACCAAGATGTCTCTTCACTAGTCATTTAGAACATGTGTAT	3072
Db	1021	CAATTTGTTAAAAGCAAGAAACCAAGATGTCTCTTCACTAGTCATTTAGAACATGTGTAT	1080
QY	3073	CATCCAAAGACTCTCTACCTCGCAACATTTGAATCCCAAGACAAATCCACATTCCTCTT	3132
Db	1081	CATCCAAAGACTCTCTACCTCGCAACATTTGAATCCCAAGACAAATCCACATTCCTCTT	1140
QY	3133	GAGTTCGCAAGCTTCTGTGTAATAGAGGCAAGCTGTCTCTATGCGCGTAAATCATCATAT	3192
Db	1141	GAGTTCGCAAGCTTCTGTGTAATAGAGGCAAGCTGTCTCTATGCGCGTAAATCATCATAT	1200
QY	3193	CTGAGGACCATTCATGGAAGCTGCTAAATAGCCTATGTGGGAGTCTTGCAATAAGTTT	3252
Db	1201	CTGAGGACCATTCATGGAAGCTGCTAAATAGCCTATGTGGGAGTCTTGCAATAAGTTT	1260
QY	3253	TGCATGAGAGCAAAACAAACAGGATTTAACTAGGTTGGTTCCCTCAGCCCTCTAAACAT	3312
Db	1261	TGCATGAGAGCAAAACAAACAGGATTTAACTAGGTTGGTTCCCTCAGCCCTCTAAACAT	1320
QY	3313	AGGGCTTAAGCTGAGGCTCTCTGGGCTTCTCTGTGTCGTAGTATTTGTAACACTAT	3372
Db	1321	AGGGCTTAAGCTGAGGCTCTCTGGGCTTCTCTGTGTCGTAGTATTTGTAACACTAT	1380
QY	3373	AGCATCTGTTAAGAATCCAGTGTCCATGGAACCTTCCCAATGCGCGTACTCTGACTAT	3432
Db	1381	AGCATCTGTTAAGAATCCAGTGTCCATGGAACCTTCCCAATGCGCGTACTCTGACTAT	1440
QY	3433	ATCATGTTTGGAAAGAGGGTTCCTGTGCTGTACAAAGCCCAAGTGAACCAAGTGTGA	3492
Db	1441	ATCATGTTTGGAAAGAGGGTTCCTGTGCTGTACAAAGCCCAAGTGAACCAAGTGTGA	1500
QY	3493	ATGCTTTTCCCTTAACACTATGTTTTTAAGTAGTCAAACTTCAGAAACATCTAAACAA	3552
Db	1501	ATGCTTTTCCCTTAACACTATGTTTTTAAGTAGTCAAACTTCAGAAACATCTAAACAA	1560
QY	3553	GTTTCTGTGGAATATGTTTGTGGAATTGTAATTGTAATTAGAGGCTCTATATTTGCA	3612
Db	1561	GTTTCTGTGGAATATGTTTGTGGAATTGTAATTGTAATTAGAGGCTCTATATTTGCA	1620
QY	3613	TTTAACTTGTTTTGTAACTCTGATCTTCTTTTGGAACTAACTATGATGAATAAAGAA	3672

```

Db      1621 TTTAACTGTTTTGTAACCTCTGATCTCTCTTTCCGATATCTATGATGAATAAGAA 1680
Qy      3673 ATTAAGTGA 3682
Db      1681 ATTAAGTGA 1690

RESULT 8
US-10-242-515-3903
; Sequence 3903, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3903
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-242-515-3903

Query Match      45.4%; Score 1678.4; DB 15; Length 1755;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1682; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      1993 AAACATTAATTTGTTTTATTTATTTCTGAGAGTTGAGGCAAAATCAGTATTTCCCAAGGAT 2052
Db      1 ANANTTTTTGTTTTATTTATTTCTGAGAGTTGAGGCAAAATCAGTATTTCCCAAGGAT 60

Qy      2053 GCGCAAGGGCAGCAAGCAGGGCTTAGGATATCCAGGCTTCAATATGCTCATTCGA 2112
Db      61 GCGCAAGGGCAGCAAGCAGGGCTTAGGATATCCAGGCTTCAATATGCTCATTCGA 120

Qy      2113 CTAACCTAGAGGGTGAAGTTGGCCCTGCTCTCTTTTCTGGACCTCAGATTTCTCAGT 2172
Db      121 CTAACCTAGAGGGTGAAGTTGGCCCTGCTCTCTTTTCTGGACCTCAGATTTCTCAGT 180

Qy      2173 GAGCTGGTAAGAAATGCATTAACCTTTGATTTGATTAAGTTAAATCTGTGGTTCGAT 2232

```

```

Db      181 GAGCTGGTAAGAAATGCATTAACCTTTTGATTTGATTAAGTTAAATCTGTGGTTCGAT 240
Qy      2233 CATTGGTCCAGAGGGAGATAGGTTCCCTGTGATTTTCTCTCTCTATAGAAATAAATG 2292
Db      241 CATTGGTCCAGAGGGAGATAGGTTCCCTGTGATTTTCTCTCTCTATAGAAATAAATG 300
Qy      2293 AAATCTGTTACTAGAAACAAGAAATGTGAGATGGCCAAAACCAAGATGACCAAGATTTGAT 2352
Db      301 AAATCTGTTACTAGAAACAAGAAATGTGAGATGGCCAAAACCAAGATGACCAAGATTTGAT 360
Qy      2353 CTCAGCTGATGACCTTACAGTCTGTGATGATGAGTCCCTCATGGTAAAGCAGGA 2412
Db      361 CTCAGCTGATGACCTTACAGTCTGTGATGATGAGTCCCTCATGGTAAAGCAGGA 420
Qy      2413 AGAGAGTGGGAAAGAAACCCACCTCTGTCTTTCATATTTGCAATTTTCATGTTTAACT 2472
Db      421 AGAGAGTGGGAAAGAAACCCACCTCTGTCTTTCATATTTGCAATTTTCATGTTTAACT 480
Qy      2473 CCGGCTGGAAATAGAAAGCAATTCCTTAGAGATGAGGATAAAGAAAGTTTCAGATTCGA 2532
Db      481 CCGGCTGGAAATAGAAAGCAATTCCTTAGAGATGAGGATAAAGAAAGTTTCAGATTCGA 540
Qy      2533 CAGGGGGAAGAAATGGAGATTTAATCTTAAACCTGTGACTTGGGAGGTCAAGTCATTTA 2592
Db      541 CAGGGGGAAGAAATGGAGATTTAATCTTAAACCTGTGACTTGGGAGGTCAAGTCATTTA 600
Qy      2593 CAGTTAGTCTGTGCTTTCGACTTCTGTGATTTAAACCCACCTCACTACCCCTGTTCA 2652
Db      601 CAGTTAGTCTGTGCTTTCGACTTCTGTGATTTAAACCCACCTCACTACCCCTGTTCA 660
Qy      2653 GATGCAATTTGGAATACCAAGATTTAAATCTCTTGACATAAGATTCATTTGCAGAAAGCAG 2712
Db      661 GATGCAATTTGGAATACCAAGATTTAAATCTCTTGACATAAGATTCATTTGCAGAAAGCAG 720
Qy      2713 ATTAAGACCATCAGAAAGAAATTTAGGTTGTAATGCACAGCACTGTGAGAACT 2772
Db      721 ATTAAGACCATCAGAAAGAAATTTAGGTTGTAATGCACAGCACTGTGAGAACT 780
Qy      2773 GTTGTCGCAAAAATAGAAATTCCTCTAGTTTTCTTTGTTTCTCATTTGAAAGGAGAAAT 2832
Db      781 GTTGTCGCAAAAATAGAAATTCCTCTAGTTTTCTTTGTTTCTCATTTGAAAGGAGAAAT 840
Qy      2833 CCACCTTTGTTAGCAATTCAGCTTTTATGATTCATCCCATCTTAAACCTCTTCAACT 2892
Db      841 CCACCTTTGTTAGCAATTCAGCTTTTATGATTCATCCCATCTTAAACCTCTTCAACT 900
Qy      2893 CCACCTTTGTTAGCAATTCAGCTTTTATGATTCATCCCATCTTAAACCTCTTCAACT 2952
Db      901 CCACCTTTGTTAGCAATTCAGCTTTTATGATTCATCCCATCTTAAACCTCTTCAACT 960
Qy      2953 AGCCTTAATCAAGGTTTTTACAGCCCTTGGACACTATGAGGAGGGGCAAGAGTACAC 3012
Db      961 AGCCTTAATCAAGGTTTTTACAGCCCTTGGACACTATGAGGAGGGGCAAGAGTACAC 1020
Qy      3013 CAATTTGTTAAAGCAAGAAACCAAGTGTCTTCTCAGTATTCATTTAGAAATGATGTTAT 3072
Db      1021 CAATTTGTTAAAGCAAGAAACCAAGTGTCTTCTCAGTATTCATTTAGAAATGATGTTAT 1080
Qy      3073 CATCCAGACTACTCTACCCCTGCAACATTTGAATCCCAAGAGCAAAATCCCAATTCCTCT 3132
Db      1081 CATCCAGACTACTCTACCCCTGCAACATTTGAATCCCAAGAGCAAAATCCCAATTCCTCTCT 1140
Qy      3133 GAGTCTGAGCTTCTGTGTAATAGGGCAGCTGTCTATGCGGTAGAAATCAGATGAT 3192
Db      1141 GAGTCTGAGCTTCTGTGTAATAGGGCAGCTGTCTATGCGGTAGAAATCAGATGAT 1200
Qy      3193 CTGAGGACCATTCATGGAAGCTGTAAATAGCTAGTCTGGGGAGTCTTCCATAAAGTTT 3252
Db      1201 CTGAGGACCATTCATGGAAGCTGTAAATAGCTAGTCTGGGGAGTCTTCCATAAAGTTT 1260
Qy      3253 TGATGAGCAAAACAAACAGGATTAACCTAGTGTGGTTTCTTCCAGCCCTCTAAAGCAT 3312
Db      1261 TGATGAGCAAAACAAACAGGATTAACCTAGTGTGGTTTCTTCCAGCCCTCTAAAGCAT 1320

```

QY 3313 AGGAGCTTACGCTGAGGCTTCTGTTGGGCTTCTCTGTGTGTAGTATTTGTAACACTAT 3372  
DB 1321 AGGGCTTAGCTGAGGCTTCTGTTGGGCTTCTCTGTGTGTAGTATTTGTAACACTAT 1380  
QY 3373 AGCATCTGTTAAAGATCCAGTGTCCATGAGAAACCTTCCACAGCCGCTGACTCTGACAT 3432  
DB 1381 AGCATCTGTTAAAGATCCAGTGTCCATGAGAAACCTTCCACAGCCGCTGACTCTGACAT 1440  
QY 3433 ATCAGTTTGGAAAGAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3492  
DB 1441 ATCAGTTTGGAAAGAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
QY 3493 ATGCTTTCTTTTACACCTATGTTTTTAAAGTGTCAAACTTCAAGAAACATCTAAACA 3552  
DB 1501 ATGCTTTCTTTTACACCTATGTTTTTAAAGTGTCAAACTTCAAGAAACATCTAAACA 1560  
QY 3553 GTTCTGTTGATATGTTGTGAACTTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 3612  
DB 1561 GTTCTGTTGATATGTTGTGAACTTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1620  
QY 3613 TTAACTGTTTGTGTAACCTGATCTTCTTTCGATTAATGATTAAGTAAGTAAGTAAGTAAGTA 3672  
DB 1621 TTAACTGTTTGTGTAACCTGATCTTCTTTCGATTAATGATTAAGTAAGTAAGTAAGTAAGTA 1680  
QY 3673 ATTAAGTGA 3682  
DB 1681 ATTAAGTGA 1690

## RESULT 9

US-09-738-973-540  
Sequence 540, Application US/09738973  
Patent No. US20020110563A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Fling, Steven P.  
APPLICANT: Mohamath, Raedoh  
APPLICANT: Algate, Paul A.  
APPLICANT: Secrist, Heather  
APPLICANT: Indictas, Carol Yoseph  
APPLICANT: Benson, Darin R.  
APPLICANT: Elliot, Mark  
APPLICANT: Mammon, Jane  
APPLICANT: Kalos, Michael D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.475C9  
CURRENT APPLICATION NUMBER: US/09/738,973  
CURRENT FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 587  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 540  
LENGTH: 634  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(634)  
OTHER INFORMATION: n = A,T,C or G  
US-09-738-973-540

Query Match 17.1%; Score 630.4; DB 9; Length 634;  
Best Local Similarity 99.5%; Pred. No. 1.1e-156;  
Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2327 CCAAAAAGAGATGACGATTTGATCTCAGCTGATGACCCCTACAGGTGCGATATGAT 2386  
DB 1 CCAAAAAGAGATGACGATTTGATCTCAGCTGATGACCCCTACAGGTGCGATATGAT 60  
QY 2387 ATGAGTCTCATGGGTAAAGCAGGAAGAGTGGAAAGAAAGAACCAACCCACTGTCT 2446

DB 61 ATGAGTCTCATGGGTAAAGCAGGAAGAGTGGAAAGAAAGAACCAACCCACTGTCT 120  
QY 2447 TCATATTTGATTTCAATGTTTAACTCCGCTGGAATTAAGATATCCCTAGAGATG 2506  
DB 121 TCATATTTGATTTCAATGTTTAACTCCGCTGGAATTAAGATATCCCTAGAGATG 180  
QY 2507 AGATTAAGAAAGTTTCAGATTCACAGGGGAGAGAAATGAGATTTAATCTTAAC 2566  
DB 181 AGATTAAGAAAGTTTCAGATTCACAGGGGAGAGAAATGAGATTTAATCTTAAC 240  
QY 2567 TGTACCTTTGGGAGGTAGTATTTAAGTATGCTGTCTTTTGAATCTTCTGATTA 2626  
DB 241 TGTACCTTTGGGAGGTAGTATTTAAGTATGCTGTCTTTTGAATCTTCTGATTA 300  
QY 2627 TTAACCCCACTACATCCCTGTTTCAAGATGATTTGAAATACCAAGATTAATCTTTGA 2686  
DB 301 TTAACCCCACTACATCCCTGTTTCAAGATGATTTGAAATACCAAGATTAATCTTTGA 360  
QY 2687 CATAGATCTCATTTGCAAGAAAGCATTTAAAGCATTCAGAAAGAAATTTAGGTTG 2746  
DB 361 CATAGATCTCATTTGCAAGAAAGCATTTAAAGCATTCAGAAAGAAATTTAGGTTG 420  
QY 2747 TATGCAAGGCAACTGTGAGAACTGTGTGCAAAATATGATTCCTTCTAGTTTTC 2806  
DB 421 TATGCAAGGCAACTGTGAGAACTGTGTGCAAAATATGATTCCTTCTAGTTTTC 480  
QY 2807 TTGTTCTATTGGAAGAGAAATTCACCTTTGTTTGAATTCAGCTTTATATGATC 2866  
DB 481 TTGTTCTATTGGAAGAGAAATTCACCTTTGTTTGAATTCAGCTTTATATGATC 540  
QY 2867 CATCCGATCTAATAAATCTTCAAACTGCTGATGATGATGATGATGATGATGATGATGATGATG 2926  
DB 541 CATCCGATCTAATAAATCTTCAAACTGCTGATGATGATGATGATGATGATGATGATGATGATG 600  
QY 2927 AAGTCCCTGAGAACTCAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2960  
DB 601 AAGTCCCTGAGAACTCAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 634

## RESULT 10

US-09-854-133-540  
Sequence 540, Application US/09854133  
Publication No. US20020183499A1  
GENERAL INFORMATION:  
APPLICANT: Lodes, Michael J.  
APPLICANT: Mohamath, Raedoh  
APPLICANT: Henderson, Robert A.  
APPLICANT: Benson, Darin R.  
APPLICANT: Secrist, Heather  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.475C10  
CURRENT APPLICATION NUMBER: US/09/854,133  
CURRENT FILING DATE: 2001-05-11  
NUMBER OF SEQ ID NOS: 735  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 540  
LENGTH: 634  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(634)  
OTHER INFORMATION: n = A,T,C or G  
US-09-854-133-540

Query Match 17.1%; Score 630.4; DB 9; Length 634;  
Best Local Similarity 99.5%; Pred. No. 1.1e-156;  
Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2327 CCAAAAAGAGATGACGATTTGATCTCAGCTGATGACCCCTACAGGTGCGATATGAT 2386

Db 1 CCAAAAACAAGATGACCCAGATTTGNTTTNAGCCTGATGACCTACAGGTCGTCTATGAT 60  
 2y 2387 ATGAGTCTCTCATGGTAAAGCAGGAGAGGTGGGAAAGAGAGACCAACCACTCTGTCT 2446  
 Db 61 ATGAGTCTCTCATGGTAAAGCAGGAGAGGTGGGAAAGAGAGACCAACCACTCTGTCT 120  
 2y 2447 TCATATTTGCAATTTTCATGTTTAACTCCGGTGGAAATAGAAAGCATTCCTCTAGAGATG 2506  
 Db 121 TCATATTTGCAATTTTCATGTTTAACTCCGGTGGAAATAGAAAGCATTCCTCTAGAGATG 180  
 2y 2507 AGGATAAAAGAAAGTTTCAGATTCAACAGGGGAGAGAAATGGAGATTAAATCCTAAAC 2566  
 Db 181 AGGATAAAAGAAAGTTTCAGATTCAACAGGGGAGAGAAATGGAGATTAAATCCTAAAC 240  
 2y 2567 TGTGACTTGGGAGGTGAGTCAATTTACAGATTAGTCTGTCTTTCGACTTCTGTGATTA 2626  
 Db 241 TGTGACTTGGGAGGTGAGTCAATTTACAGATTAGTCTGTCTTTCGACTTCTGTGATTA 300  
 2y 2627 TTAACCCCACTCACTACCTGTTTCAGATGATTTGGAATACCAAGATTAAATCCTTGA 2686  
 Db 301 TTAACCCCACTCACTACCTGTTTCAGATGATTTGGAATACCAAGATTAAATCCTTGA 360  
 2y 2687 CATAGATCTCATTTGAGAAAGCAGATTAAAGACCATCAAGAGAAATTAATTTAGTTG 2746  
 Db 361 CATAGATCTCATTTGAGAAAGCAGATTAAAGACCATCAAGAGAAATTAATTTAGTTG 420  
 2y 2747 TAATGCACAGCAACTGTGAGAAACTGTGTGCGCAAAATAGAAATTCCTCTAGTTTTC 2806  
 Db 421 TAATGCACAGCAACTGTGAGAAACTGTGTGCGCAAAATAGAAATTCCTCTAGTTTTC 480  
 2y 2807 TTGTTCTCATTTGAAAGAGAGAAATTCACATTTGTTAGCATTTCAAGCTTTTATGTATC 2866  
 Db 481 TTGTTCTCATTTGAAAGAGAGAAATTCACATTTGTTAGCATTTCAAGCTTTTATGTATC 540  
 2y 2867 CATCCCATCTAAAACTCTTCAAACTCCACTGTTTCAGTCTGAAATGCAGTCCCTGTCC 2926  
 Db 541 CATCCCATCTAAAACTCTTCAAACTCCACTGTTTCAGTCTGAAATGCAGTCCCTGTCC 600  
 2y 2927 AAGTGCCCTTGAGAACTCACAGCAGCGCCTTA 2960  
 Db 601 AAGTGCCCTTGAGAACTCACAGCAGCGCCTTA 634

RESULT 11

US-10-144-649A-540  
 ; Sequence 540, Application US/10144649A  
 ; Publication No. US20030118599A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Fan, Liqun  
 ; APPLICANT: Algate, Paul A.  
 ; APPLICANT: McNeill, Patricia D.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
 ; THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.475C11  
 ; CURRENT APPLICATION NUMBER: US/10/144,649A  
 ; CURRENT FILING DATE: 2002-08-21  
 ; NUMBER OF SEQ ID NOS: 749  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 540  
 ; LENGTH: 634  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: 25..79  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-10-144-649A-540

Query Match 17.1%; Score 630.4; DB 14; Length 634;  
 Best Local Similarity 99.5%; Pred. No. 1.1e-156;  
 Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2327 CCAAAAACAAGATGACCCAGATTTGATCTCAGCTGATGACCTACAGTCTGTCTATGAT 2386  
 Db 1 CCAAAAACAAGATGACCCAGATTTGATCTCAGCTGATGACCTACAGTCTGTCTATGAT 60  
 QY 2387 ATGAGTCTCTCATGGTAAAGCAGGAGAGGTGGGAAAGAGAGACCAACCACTCTGTCT 2446  
 Db 61 ATGAGTCTCTCATGGTAAAGCAGGAGAGGTGGGAAAGAGAGACCAACCACTCTGTCT 120  
 QY 2447 TCATATTTGCAATTTTCATGTTTAACTCCGGTGGAAATAGAAAGCATTCCTCTAGAGATG 2506  
 Db 121 TCATATTTGCAATTTTCATGTTTAACTCCGGTGGAAATAGAAAGCATTCCTCTAGAGATG 180  
 QY 2507 AGGATAAAAGAAAGTTTCAGATTCAACAGGGGAGAGAAATGGAGATTAAATCCTAAAC 2566  
 Db 181 AGGATAAAAGAAAGTTTCAGATTCAACAGGGGAGAGAAATGGAGATTAAATCCTAAAC 240  
 QY 2567 TGTGACTTGGGAGGTGAGTCAATTTACAGATTAGTCTGTCTTTCGACTTCTGTGATTA 2626  
 Db 241 TGTGACTTGGGAGGTGAGTCAATTTACAGATTAGTCTGTCTTTCGACTTCTGTGATTA 300  
 QY 2627 TTAACCCCACTCACTACCTGTTTCAGATGATTTGGAATACCAAGATTAAATCCTTGA 2686  
 Db 301 TTAACCCCACTCACTACCTGTTTCAGATGATTTGGAATACCAAGATTAAATCCTTGA 360  
 QY 2687 CATAGATCTCATTTGAGAAAGCAGATTAAAGACCATCAAGAGAAATTAATTTAGTTG 2746  
 Db 361 CATAGATCTCATTTGAGAAAGCAGATTAAAGACCATCAAGAGAAATTAATTTAGTTG 420  
 QY 2747 TAATGCACAGCAACTGTGAGAAACTGTGTGCGCAAAATAGAAATTCCTCTAGTTTTC 2806  
 Db 421 TAATGCACAGCAACTGTGAGAAACTGTGTGCGCAAAATAGAAATTCCTCTAGTTTTC 480  
 QY 2807 TTGTTCTCATTTGAAAGAGAGAAATTCACATTTGTTAGCATTTCAAGCTTTTATGTATC 2866  
 Db 481 TTGTTCTCATTTGAAAGAGAGAAATTCACATTTGTTAGCATTTCAAGCTTTTATGTATC 540  
 QY 2867 CATCCCATCTAAAACTCTTCAAACTCCACTGTTTCAGTCTGAAATGCAGTCCCTGTCC 2926  
 Db 541 CATCCCATCTAAAACTCTTCAAACTCCACTGTTTCAGTCTGAAATGCAGTCCCTGTCC 600  
 QY 2927 AAGTGCCCTTGAGAACTCACAGCAGCGCCTTA 2960  
 Db 601 AAGTGCCCTTGAGAACTCACAGCAGCGCCTTA 634

RESULT 12

US-09-735-705-209  
 ; Sequence 209, Application US/09735705  
 ; Patent No. US20020052329A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Fan, Liqun  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Bangur, Chaitanya S.  
 ; APPLICANT: Hosken, Nancy R.  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Li, Samuel X.  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Fanger, Neil  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; AND DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.455C14  
 ; CURRENT APPLICATION NUMBER: US/09/735,705  
 ; CURRENT FILING DATE: 2000-12-12  
 ; NUMBER OF SEQ ID NOS: 419  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 209  
 ; LENGTH: 621  
 ; TYPE: DNA

ORGANISM: Homo sapiens  
US-09-735-703-209

Query Match 16.8%; Score 619.4; DB 9; Length 621;  
Best Local Similarity 99.8%; Pred. No. 9.1e-154;  
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3055 CATTAGAACATGGTTATCATCCAGACTCTGACCCCTGCAACATTGAACTGCCAAG 3114  
DB 1 CATTAGAACATGGTTATCATCCAGACTCTGACCCCTGCAACATTGAACTGCCAAG 60  
QY 3115 CAATCCACATTCCTCTTGAAGTTCTGACGCTCTGTGTAAATAGGCGAGCTGTCTAT 3174  
DB 61 CAATCCACATTCCTCTTGAAGTTCTGACGCTCTGTGTAAATAGGCGAGCTGTCTAT 120  
QY 3175 GCCGTAAATCAGTATCTGAGAGCAATTCATGGAAGCTGCTAAATAGCTGTGG 3234  
DB 121 GCCGTAAATCAGTATCTGAGAGCAATTCATGGAAGCTGCTAAATAGCTGTGG 180  
QY 3235 GAGTCTTCATTAAGTTTGGATGAGCAAAACAAGATTAACCTAGCTTGTCT 3294  
DB 181 GAGTCTTCATTAAGTTTGGATGAGCAAAACAAGATTAACCTAGCTTGTCT 240  
QY 3295 TCAGCCCTCTAAAGCATAGGCTTAGCTGCAAGCTTCTTGGGCTTCTGTGTGTG 3354  
DB 241 TCAGCCCTCTAAAGCATAGGCTTAGCTGCAAGCTTCTTGGGCTTCTGTGTGTG 300  
QY 3355 TAGTTTGTAAACATATAGCATCTGTTAAGATCCAGTTCATGGAACCTTCCACAT 3414  
DB 301 TAGTTTGTAAACATATAGCATCTGTTAAGATCCAGTTCATGGAACCTTCCACAT 360  
QY 3415 GCCGTGACTGGAATATAGTTTGGAAAGCAGGTTCTCTGCTGCTAACAGC 3474  
DB 361 GCCGTGACTGGAATATAGTTTGGAAAGCAGGTTCTCTGCTGCTAACAGC 420  
QY 3475 CCAGTGGACAGCTGATGATCTTCTCTTACACATGTTTAAAGTAGTCAAACTTC 3534  
DB 421 CCAGTGGACAGCTGATGATCTTCTCTTACACATGTTTAAAGTAGTCAAACTTC 480  
QY 3535 AAGAAACATCTAAACAAGTTCTGTGATATGTGTGAACTGTGATTTA 3594  
DB 481 AAGAAACATCTAAACAAGTTCTGTGATATGTGTGAACTGTGATTTA 540  
QY 3595 GTAGGCTCTATATGATTAATGATTTTGTAGTCCGATCTTCTTTCGATA 3654  
DB 541 GTAGGCTCTATATGATTAATGATTTTGTAGTCCGATCTTCTTTCGATA 600  
QY 3655 CTATTGATGAATAAAGAAATT 3675  
DB 601 CTATTGATGAATAAAGAAATT 621

RESULT 13  
US-09-850-716A-209

Sequence 209, Application US/09850716A  
Patent No. US20020115139A1  
GENERAL INFORMATION:  
APPLICANT: Kalo, Michael D.  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Retter, Marc W.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.455C15  
CURRENT APPLICATION NUMBER: US/09/850.716A  
CURRENT FILING DATE: 2001-05-07  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 209  
LENGTH: 621  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-850-716A-209

Query Match 16.8%; Score 619.4; DB 9; Length 621;  
Best Local Similarity 99.8%; Pred. No. 9.1e-154;  
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3055 CATTAGAACATGGTTATCATCCAGACTCTGACCCCTGCAACATTGAACTGCCAAG 3114  
DB 1 CATTAGAACATGGTTATCATCCAGACTCTGACCCCTGCAACATTGAACTGCCAAG 60  
QY 3115 CAATCCACATTCCTCTTGAAGTTCTGACGCTCTGTGTAAATAGGCGAGCTGTCTAT 3174  
DB 61 CAATCCACATTCCTCTTGAAGTTCTGACGCTCTGTGTAAATAGGCGAGCTGTCTAT 120  
QY 3175 GCCGTAAATCAGTATCTGAGAGCAATTCATGGAAGCTGCTAAATAGCTGTGG 3234  
DB 121 GCCGTAAATCAGTATCTGAGAGCAATTCATGGAAGCTGCTAAATAGCTGTGG 180  
QY 3235 GAGTCTTCATTAAGTTTGGATGAGCAAAACAAGATTAACCTAGCTTGTCT 3294  
DB 181 GAGTCTTCATTAAGTTTGGATGAGCAAAACAAGATTAACCTAGCTTGTCT 240  
QY 3295 TCAGCCCTCTAAAGCATAGGCTTAGCTGCAAGCTTCTTGGGCTTCTGTGTGTG 3354  
DB 241 TCAGCCCTCTAAAGCATAGGCTTAGCTGCAAGCTTCTTGGGCTTCTGTGTGTG 300  
QY 3355 TAGTTTGTAAACATATAGCATCTGTTAAGATCCAGTTCATGGAACCTTCCACAT 3414  
DB 301 TAGTTTGTAAACATATAGCATCTGTTAAGATCCAGTTCATGGAACCTTCCACAT 360  
QY 3415 GCCGTGACTGGAATATAGTTTGGAAAGCAGGTTCTCTGCTGCTAACAGC 3474  
DB 361 GCCGTGACTGGAATATAGTTTGGAAAGCAGGTTCTCTGCTGCTAACAGC 420  
QY 3475 CCAGTGGACAGCTGATGATCTTCTCTTACACATGTTTAAAGTAGTCAAACTTC 3534  
DB 421 CCAGTGGACAGCTGATGATCTTCTCTTACACATGTTTAAAGTAGTCAAACTTC 480  
QY 3535 AAGAAACATCTAAACAAGTTCTGTGATATGTGTGAACTGTGATTTA 3594  
DB 481 AAGAAACATCTAAACAAGTTCTGTGATATGTGTGAACTGTGATTTA 540  
QY 3595 GTAGGCTCTATATGATTAATGATTTTGTAGTCCGATCTTCTTTCGATA 3654  
DB 541 GTAGGCTCTATATGATTAATGATTTTGTAGTCCGATCTTCTTTCGATA 600  
QY 3655 CTATTGATGAATAAAGAAATT 3675  
DB 601 CTATTGATGAATAAAGAAATT 621

RESULT 14  
US-09-897-778-209

Sequence 209, Application US/09897778  
Patent No. US20020147143A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Tonglong  
APPLICANT: Wainerakis, Margarita  
APPLICANT: Fanger, Gary R.  
APPLICANT: Vedvik, Thomas S.  
APPLICANT: Carter, Derrick  
APPLICANT: Matanabe, Yoshihiro  
APPLICANT: Henderson, Robert A.  
APPLICANT: Beckham, David W.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.455C16  
CURRENT APPLICATION NUMBER: US/09/897.778  
CURRENT FILING DATE: 2001-06-28  
NUMBER OF SEQ ID NOS: 467  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 209  
LENGTH: 621  
TYPE: DNA

ORGANISM: Homo sapiens  
JS-09-897-778-209

Query Match 16.8%; Score 619.4; DB 9; Length 621;  
Best Local Similarity 99.8%; Pred. No. 9.1e-154;  
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2y 3055 CATTAGAACATGGTTATCATCCAGACTACTCTACCTGCAACATTGAACCTCCAGAG 3114  
Db 1 CATTAGAACATGGTTATCATCCAGACTACTCTACCTGCAACATTGAACCTCCAGAG 60

2y 3115 CAAATCCACATTCCTCTTGGTCTGAGCTTCTGTGTAAATAGGCGAGCTGTGCTAT 3174  
Db 61 CAAATCCACATTCCTCTTGGTCTGAGCTTCTGTGTAAATAGGCGAGCTGTGCTAT 120

2y 3175 GCCGTAGAACATCATGATCTGAGACCAATTCATGGAAGCTGCTAAATAGCTAGTCTGG 3234  
Db 121 GCCGTAGAACATCATGATCTGAGACCAATTCATGGAAGCTGCTAAATAGCTAGTCTGG 180

2y 3235 GAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTAACTAGGTTGGTTCCT 3294  
Db 181 GAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTAACTAGGTTGGTTCCT 240

2y 3295 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGCTTCCCTGGGCTTCTCTGTGTG 3354  
Db 241 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGCTTCCCTGGGCTTCTCTGTGTG 300

2y 3355 TAGTTTTGTAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACAT 3414  
Db 301 TAGTTTTGTAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACAT 360

2y 3415 GCCGTGACTCTGGAATATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTAACAGC 3474  
Db 361 GCCGTGACTCTGGAATATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTAACAGC 420

2y 3475 CCAGTGGACCAAGTCTGATGCTTTCCTTTTACACCTATGTTTAAAGTAGTCAAACTTC 3534  
Db 421 CCAGTGGACCAAGTCTGATGCTTTCCTTTTACACCTATGTTTAAAGTAGTCAAACTTC 480

2y 3535 AAGAAACAATCTAAACAAGTTTCTGTCATATGTTTGTGAACCTTGTATTGTAATTA 3594  
Db 481 AAGAAACAATCTAAACAAGTTTCTGTCATATGTTTGTGAACCTTGTATTGTAATTA 540

2y 3595 GTAGGCTTCTATATGCAATTAACCTGTTTGTAACTCCTGATTCCTCTTCGATA 3654  
Db 541 GTAGGCTTCTATATGCAATTAACCTGTTTGTAACTCCTGATTCCTCTTCGATA 600

2y 3655 CTATTGATGAATAAAGAAATT 3675  
Db 601 CTATTGATGAATAAAGAAATT 621

RESULT 15  
US-09-466-396A-209  
Sequence 209, Application US/09466396A  
Publication No. US20030119763A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND  
FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.455C4  
CURRENT APPLICATION NUMBER: US/09/466.396A  
CURRENT FILING DATE: 1999-12-17  
NUMBER OF SEQ ID NOS: 224  
SOFTWARE: Fast-Seq for Windows Version 3.0  
SEQ ID NO 209  
LENGTH: 621  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-466-396A-209

Query Match 16.8%; Score 619.4; DB 10; Length 621;  
Best Local Similarity 99.8%; Pred. No. 9.1e-154;

Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3055 CATTAGAACATGGTTATCATCCAGACTACTCTACCTGCAACATTGAACCTCCAGAG 3114  
Db 1 CATTAGAACATGGTTATCATCCAGACTACTCTACCTGCAACATTGAACCTCCAGAG 60

QY 3115 CAAATCCACATTCCTCTTGGTCTGAGCTTCTGTGTAAATAGGCGAGCTGTGCTAT 3174  
Db 61 CAAATCCACATTCCTCTTGGTCTGAGCTTCTGTGTAAATAGGCGAGCTGTGCTAT 120

QY 3175 GCCGTAGAACATCATGATCTGAGACCAATTCATGGAAGCTGCTAAATAGCTAGTCTGG 3234  
Db 121 GCCGTAGAACATCATGATCTGAGACCAATTCATGGAAGCTGCTAAATAGCTAGTCTGG 180

QY 3235 GAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTAACTAGGTTGGTTCCT 3294  
Db 181 GAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTAACTAGGTTGGTTCCT 240

QY 3295 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGCTTCCCTGGGCTTCTCTGTGTG 3354  
Db 241 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGCTTCCCTGGGCTTCTCTGTGTG 300

QY 3355 TAGTTTTGTAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACAT 3414  
Db 301 TAGTTTTGTAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACAT 360

QY 3415 GCCGTGACTCTGGAATATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTAACAGC 3474  
Db 361 GCCGTGACTCTGGAATATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTAACAGC 420

QY 3475 CCAGTGGACCAAGTCTGATGCTTTCCTTTTACACCTATGTTTAAAGTAGTCAAACTTC 3534  
Db 421 CCAGTGGACCAAGTCTGATGCTTTCCTTTTACACCTATGTTTAAAGTAGTCAAACTTC 480

QY 3535 AAGAAACAATCTAAACAAGTTTCTGTCATATGTTTGTGAACCTTGTATTGTAATTA 3594  
Db 481 AAGAAACAATCTAAACAAGTTTCTGTCATATGTTTGTGAACCTTGTATTGTAATTA 540

QY 3595 GTAGGCTTCTATATGCAATTAACCTGTTTGTAACTCCTGATTCCTCTTCGATA 3654  
Db 541 GTAGGCTTCTATATGCAATTAACCTGTTTGTAACTCCTGATTCCTCTTCGATA 600

QY 3655 CTATTGATGAATAAAGAAATT 3675  
Db 601 CTATTGATGAATAAAGAAATT 621

Search completed: March 12, 2004, 06:13:28  
Job time : 865.651 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:18:38 ; Search time 6114.16 seconds  
(without alignments)  
18036.975 Million cell updates/sec

Title: US-10-084-817-20  
Perfect score: 3693  
Sequence: 1 cgaatcgagactccaccct.....ttaaagtgaataaaaaaa 3693

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:
- 1: em\_estba:\*\*
  - 2: em\_esthum:\*\*
  - 3: em\_estin:\*\*
  - 4: em\_estmu:\*\*
  - 5: em\_estov:\*\*
  - 6: em\_estpl:\*\*
  - 7: em\_estro:\*\*
  - 8: em\_htc:\*\*
  - 9: gb\_est1:\*\*
  - 10: gb\_est2:\*\*
  - 11: gb\_htc:\*\*
  - 12: gb\_est3:\*\*
  - 13: gb\_est4:\*\*
  - 14: gb\_est5:\*\*
  - 15: em\_estfun:\*\*
  - 16: em\_estom:\*\*
  - 17: em\_gss\_hum:\*\*
  - 18: em\_gss\_inv:\*\*
  - 19: em\_gss\_pln:\*\*
  - 20: em\_gss\_vrt:\*\*
  - 21: em\_gss\_fun:\*\*
  - 22: em\_gss\_mam:\*\*
  - 23: em\_gss\_mus:\*\*
  - 24: em\_gss\_pro:\*\*
  - 25: em\_gss\_rod:\*\*
  - 26: em\_gss\_phg:\*\*
  - 27: em\_gss\_vrl:\*\*
  - 28: gb\_gss1:\*\*
  - 29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1050	28.4	3610	11 AK049440	AK049440 Mus muscu
2	1044.6	28.3	1110	9 AL548992	AL548992 AL548992
3	1007.6	27.3	1109	9 AL573353	AL573353 AL573353
4	991.2	26.8	1090	12 BM553173	BM553173 AGENCOURT

5	988.6	26.8	1094	12 BM463192	BM463192 AGENCOURT
6	960.4	26.0	1677	11 BC006185	BC006185 Mus muscu
7	913.6	24.7	999	9 AL570818	AL570818 AL570818
8	905.8	24.5	1048	13 BX338760	BX338760 BX338760
9	905.6	24.5	1201	13 BX405684	BX405684 BX405684
10	879.6	23.8	995	9 AL544661	AL544661 AL544661
11	875.6	23.7	987	13 BX395667	BX395667 BX395667
12	869.6	23.5	1201	13 BX352901	BX352901 BX352901
13	845.8	22.9	916	13 BQ671550	BQ671550 AGENCOURT
14	844	22.9	879	9 AB045227	AB045227 AB045227
15	844	22.9	927	13 BQ918886	BQ918886 AGENCOURT
16	842.8	22.8	906	13 BUI46207	BUI46207 AGENCOURT
17	840.8	22.8	917	13 BQ718490	BQ718490 AGENCOURT
18	838.8	22.7	891	14 CF242911	CF242911 AGENCOURT
19	835	22.6	1201	13 BX405683	BX405683 BX405683
20	825.4	22.4	1068	12 BM912228	BM912228 AGENCOURT
21	821.4	22.2	878	14 CD518823	CD518823 AGENCOURT
22	821.4	22.2	911	13 BQ883171	BQ883171 AGENCOURT
23	814.8	22.1	886	13 BQ932018	BQ932018 AGENCOURT
24	807.4	21.9	1142	12 BM916732	BM916732 AGENCOURT
25	801.2	21.7	853	13 BU517310	BU517310 AGENCOURT
26	797.2	21.6	829	9 AU140319	AU140319 AU140319
27	795.2	21.5	933	13 BQ897777	BQ897777 AGENCOURT
28	791.2	21.5	809	9 AU130022	AU130022 AU130022
29	791.2	21.4	892	13 BQ707255	BQ707255 AGENCOURT
30	789.8	21.4	831	9 AU131134	AU131134 AU131134
31	786.4	21.3	833	9 AU132046	AU132046 AU132046
32	784.4	21.2	904	14 CD244996	CD244996 AGENCOURT
33	783.2	21.2	840	12 BG570724	BG570724 602591838
34	781.4	21.2	872	14 CD515270	CD515270 AGENCOURT
35	780.6	21.1	929	13 BQ722506	BQ722506 AGENCOURT
36	778	21.1	799	14 CK000721	CK000721 AGENCOURT
37	773.6	20.9	875	14 CD388013	CD388013 AGENCOURT
38	772.6	20.9	964	13 BQ707449	BQ707449 AGENCOURT
39	771.4	20.9	1030	13 BQ952029	BQ952029 AGENCOURT
40	771.2	20.9	910	13 BQ719949	BQ719949 AGENCOURT
41	769.8	20.8	876	14 CD654748	CD654748 AGENCOURT
42	767	20.8	885	9 AL710894	AL710894 DKFZ9686G
43	766.6	20.8	801	14 CK000499	CK000499 AGENCOURT
44	766.4	20.8	1101	9 AL574665	AL574665 AL574665
45	756.2	20.5	773	14 CA414235	CA414235 UI-H-E20-

ALIGNMENTS

RESULT 1	AK049440	3610 bp	mRNA	linear	HTC 20-SEP-2003
LOCUS	Mus musculus	7 days embryo whole body cDNA, RIKEN full-length			
DEFINITION	enriched library, clone: C43001K03	product: follistatin-like, full insert sequence.			
ACCESSION	AK049440				
VERSION	AK049440.1	GI:26340175			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	1				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				

REFERENCE	3	Shibata, K., Itch, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Komori, H., Akiyama, J., Nishikawa, K., Katsunuma, T., Teshiro, H., Itch, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, N., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, Y., Inoue, K., Togawa, Y., Izawa, M., Obara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multiplexed sequencing
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)	
MEDLINE	20530913	
PUBMED	11076861	
REFERENCE	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection	
JOURNAL	Nature 409, 685-690 (2001)	
REFERENCE	5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 50,770 full-length cDNAs	
JOURNAL	Nature 420, 563-573 (2002)	
REFERENCE	6 (bases 1 to 3610)	
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagata, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirokawa, T., Hirozane, T., Horii, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanakura, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komori, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akehata, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-Jul-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsukuba, Ibaraki, Japan 305-0856, Japan (E-mail: genome-research@riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.	
FEATURES	location/Qualifiers	
source	1. 3610	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strain="C57BL/6J"	
	/db_xref="FANTOM DB:C430011K03"	
	/db_xref="MGI:2416783"	
	/db_xref="taxon:10090"	
	/clone="C430011K03"	
	/csize_type="whole body"	
	/clone_lib="RIKEN full-length enriched mouse cDNA library"	
	/dev_stage="7 days embryo"	
	100..1020	
	/note="unmated protein product; foliostatin-like (MGI:MGI:102793, GB NM_008047, evidence: BLASTN, 99%, match=2819)	
	/codon_start=1	
	/protein_id="BAC3751.1"	
	/db_xref="GI:26340176"	
CDS		

Query Match	28.4%	Score 1050	DB 11	Length 3610
Best Local Similarity	76.9%	Pred. No. 1e-200		
Matches 1591	Conservative 0	Mismatches 390	Indels 87	Gaps 22
QY 12	CTCCACACCTCCGCTTACAGCTCGCTGCGCGCGCTCTGCCGCCGCCGCCGCCAGAGAGACTTGG			71
DB 29	CTCCACACCTTGGCTCTTAACCTGCTGCGCGCACCTGCCCAAGTGTCTCTCGGAGTCCCGG			88
QY 72	ACCCAGACCAACATGTGTGGAAAAGCTGTGCTCGCGCTCGCGCTCGCGCTGTGTGGGCGTCCCT			131
DB 89	ACCCAGACCAACATGTGTGGAAAAGATGG-----CTGGGCGCTCTGCTGTGGACCAATCGGCC			142
QY 132	GGGTCCGCGCCGAGGAGAGCTAAGAGCAAAATCCAAAGATCTGTGCCAAATGTGTGTTTGTG			191
DB 143	TGGTCCAGGGGAGGAGGAACTTGAAAGCAAAATCCAAAGATCTGCGCAATGTGTGTTTGTG			202
QY 192	GAGCGCGCGCGGAAATGTGTGACATGCACAGAAAAGGGGAAACCCACTGTCTCTGCATTTGAGC			251
DB 203	GAGCTGGCAGGGAATGTGTCCGTGCACAGAGAAAGGGGAGCCCAAGTGTCTCTGTGATTGAGC			262
QY 252	AATGCAAACTTCACAAGAGGCGCTGTGTGTGTGGCAGTATGTGACAAGAACCTTACCTCAACCT			311
DB 263	AATGCAAACTTCACAAGAGGCGCTGTGTGTGTGGCAGTATGTGACAAGAACCTTACCTCAACCT			322
QY 312	GTGAACTGATGAGAAATGCTGTGCTCTCACTGGATTCGAAAATCCGAGGTGTGATTAGAGAGAC			371
DB 323	GTGAACCTTATGAAATGCTGTGCTCTCACTGGATTCGAAAATCCGAGGTGTGATTAGAGAGAC			382
QY 372	ACTGCAAAAGAGAAATCCGTAAATGTCATCTGCGACGCCAGTGTGTTGTCTATTCAGTCCA			431
DB 383	ACTGCAAAAGAAAGATCTGCAAGTCCATCTGCGACGCCAGTGTGTTGTCTATTCAGGCTA			442
QY 432	ACCGATATGAGCTCCGAGTGTGCATATCCAGTGTGCGTGAAGGCTGTGAGATCATCTGAGATG			491
DB 443	ACCGCATATGAGCTCCGAGTGTGCATATCCAGTGTGCGTGAAGGCTGTGAGATCATCTGAGATG			502
QY 492	GCTGTGTTCTTAAAGGCGCAACTACAGTGAATCTCTGACCAAGTATTTTAAAGACTTGG			551
DB 503	GCTGTGTTCTTAAAGGCGAGTAACTACAGTGAATCTCTGACCAAGTATTTTAAAGACTTGG			562
QY 552	ATPATGTGATTTCTGGCTGTGACCTCCAGTGAATTCCTGGAAGTTTGTGAAACGAATGAAA			611
DB 563	ATPATGTGCACTCTCACTTGAATCTCCAGTGAATTCCTGGAAGTTTGTGAGACGAATGAAA			622
QY 612	CTGCACTCAATATTAACAAGTATCCAGACGAGAGAACCAAGTGTCTTAAAGGAGCTCT			671
DB 623	CAGCCATCAACATCAACCACTTATGCAATGACGAGAGAACCAAGTGTCTCAGAAGCTCT			682
QY 672	GTGTGTATGCTCTCATTTGAATGTGTGTATGAAATGTGTATGAAACTCAGTTTCAAG			731
DB 683	GTGTGTATGCTCTCATTTGAATGTGTGTATGAAAGCTGTGATGAAACTCAGTTTCAAG			742
QY 732	AGTTTCTCAAGTGCCTCAACCACTCTTCAACCTCTCTGGAAGAAAGTGTGCCCTGAGAG			791
DB 743	AGTTTCTCAAGTGCCTCAACCACTCTTCAACCTCTCTGGAAGAAAGTGTGCCCTGAGAG			802
QY 792	ATGAAACGATACAGATGAGCTGAGACCGAGGTGACCTTAAACCGCTGTGTGCT			851
DB 803	ACGAAACCTATGCAATGAGCTGAGACTGAGGTGACCTGCAATCCGTGTGTGTGTCT			862
QY 852	GTGGAATTTGGTCTGTACAGCATGACTGTGTGAGGAAAGAAATCAAGAAAGGGGCGCCAGA			911

```
|||||
863 GTGCCACTGGGTCTGCACAGCAATACCTGTGATGGAAGAATCAGAAGGGGGTCCAGA 922
|||||
912 CCAGACAGAGGAGGAGATGACAGATATGTCTCAGGAGCTCCAAAAGCATCAGGAACAG 971
|||||
923 CCACACAGAGGAGGAGAGCAGGATATGTCCAGGAATCCAGAGACACAGGGGACAG 982
|||||
972 CTGAAAAGACCAAGAGAGTGAGCACCAGAGAGATCTAATGAG-GAGGCCACAGACCAAGTGT 1030
|||||
983 CAGAAAAGACCAAGAGAGTGAAACACCAAGAGATCTAAGAGAGGACACAGACCGTGT 1042
|||||
1031 CTGATGCCAGCATCTTCTCCATCTCAGCGCTCAGTTCAGTATACACAAGTGTCTGCTAC 1090
|||||
1043 CCGAGGCCAGCGCCT---CCTCTCAGCGCTCAGCCAGTACACAC-AGAGTCTGCAGC 1098
|||||
1091 AGTCGCCAAATCACCAGTATTTCTTATPAGCAATGAGTTTTTATTTGTTTATTTGTTT 1150
|||||
1099 AATCACCAAATCACTAGTATTTCTTGTATGGCAGCAATCTTATTTTG---TTTGTTT 1154
|||||
1151 TCGAATTAAGGATATGAGTGTGCTGTAGGAGGAGGAGGCCACAGCTTCATTTCTA 1210
|||||
1155 TCGAATTAAGGATATGAGTGTGCTGTAGGAGGAGGAGGCCACAACTTCACCTGTA 1214
|||||
1211 GGAAGTCTTTAAGAGAACTGTAAATGGTCTCTGGGGCTGAGGCTAGTAAAGGAACTG 1270
|||||
1215 GGAATGCTTTAAGAGAAC--TAAAGACACCTTGGACGAGAGCAACTAAGGAAACAG 1272
|||||
1271 CATCAAGTTAAAGAGGAAACAGACCAATCTGAAC---CTCTTTGAGTTTACTGCA 1326
|||||
1273 CATC-GGGTTGGCAGGAGGACAGAGGAGGTTTGAATGAAGCCTTTCTGGGGTCACAGCA 1331
|||||
1327 TCTGTACAGAGGCTGACGAGGAGTGCAACGATGCCAGAGAGAACTTAGCAGGGTGTCCCC 1386
|||||
1332 CTGCGGAGGAGATACAGGAAGCAT-----AGGAACACTTGAAGTCCCTG 1381
|||||
1387 GAGGAGAGGTTTGGAGAGTCCACGAGAGGAAACGCTCTCTGCTTCCAGCTCTTTTCCA 1446
|||||
1382 CTGAGGAGGTTTGGGGAGGCTTTGTAGGAGGAAACCTCTGCTGTTT---GACCTTTGCA 1438
|||||
1447 TTGCGGTGACATGACAGACCTCAGCATCCAGCATCTCTGGTCCCAATAAC-TGCCT 1505
|||||
1439 CCACTGTCACTGACAGACCTGACGACAGTCTGCTTCTCTTTGGTCCCAACATCAC 1498
|||||
1506 CTAGATACATAGCATCTGCTAGTTAAACCAAGTGCCTCAGACTTGGATGAGTTCT 1565
|||||
1499 CTGAACACACAGCGGCCCAACTAGTTA--CCTGTGCTCCTCAGCCTTGCATGGAGTTCC 1556
|||||
1566 GGGAGGTACACCAATGATGAGATCTGTATCTTGTAGCCCTTACGACCTAC 1625
|||||
1557 TGGAGGAGGTTTAAATGATGACAGACTTATGTACTTCAAGCGCATGGAGACCTAAC 1616
|||||
1626 CAAATTTTAAAAATACTTTTT-----ACCAAGGTGCTATTTCT 1664
|||||
1617 CAAATTTTAAAAATACATTTTCTTTTTTTTTTTTGTGTTAACCAAGGTGCTATTTCT 1676
|||||
1665 CTGTAAACACTTTTTTGGCAGTGTACTTTTCTTCAATTTATATATATATAT 1724
|||||
1677 CTGTAAAGAGCTTTTTT---CCAAGTGTACTTTCCTCAGTATATACCGTTATATAT 1733
|||||
1725 TGT--TTTTTAAATATTTTATTTTCTGACTAGGTATTAAGCTTTTGTAAATTA-TTTTTCA 1781
|||||
1734 TGTGTTTTTTTAAATATTTTCAATTTTGTACTAGATATTAAGCTTTTGTAAATTTTTC 1793
|||||
1782 GTAGTCCACACTTCTAGTGTGA-----AGGAGTTGGGGTTCTTCTGTTGTCAGG 1834
|||||
1794 TTAGTCTCTACTATTTTCAGAGTGAAGGTGAAGGGGTTTGGGCATTTTCCAGGTACAGG 1853
|||||
1835 GGCTGAATAAACCCAGA-TGCTCCACCTTGCACATACCTAGTA-----TGACGCCCATAG 1888
|||||
1854 GAATCTGTAAACACACAGCCCATACCTCTGTACATATTTAGACCGGTTGCAGTTCGGAG 1913
|||||
1889 TTGGCC-----CCCTAGCTTTCAGCAGTCCA---CTATCTGCCAGAGGACGAGGTC 1940
|||||
```

```
Db 1914 CTGACCCCCAACCCAGAGCTTTCTAGAAAATCAGCTCCATCCACGAAGCACAAAGGCC 1973
Qy 1941 CTTAGACCTTAAGCCAGGGAGAGCATCTTTCTATAAAAAAACTTTCAAGATCCCAACA-TT 1999
Db 1974 CTCAGACAGAAGCCACAGGACAAAGCATCTTCATAGACAGCTGTGTGATCCCAACAGTT 2033
Qy 2000 AATTGTTTTTTATTTATTTCTGAGAACTT 2027
Db 2034 AATTGCTTTTGTCTTCTTGAAGAAGTT 2061

RESULT 2
AL548992
LOCUS AL548992 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1042YC07 5-PRIME, mRNA sequence.
ACCESSION AL548992
VERSION AL548992.2 GI:31270813
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1110)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12884545.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1910.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1042AB04QPI&cluster=1910.r. Contact :
Feng liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1042AB04QPI.

FEATURES
Location/Qualifiers
1..1110
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1042YC07"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 28.3%; Score 1044.6; DB 9; Length 1110;
Best Local Similarity 99.3%; Pred. No. 1.6e-199;
Matches 1044; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GATCGGCGACTCCACCTCGCTTACAGCTCGCTCGCGCGTCTCGCGCGCCCA 61
Db 60 GATCGGCGACTCCACCTCGCTTACAGCTCGCTCGCGCGTCTCGCGCGCCCA 119
Qy 62 GGAGACCTGGACACGACGATGTGGAACGCTCGGCTCGGCTCGGCTCGGCTCGG 121
Db 120 GGAGACCTGGACACGACGATGTGGAACGCTCGGCTCGGCTCGGCTCGGCT 179
Qy 122 GCGGTGCGCTCGGTCGCGCGGAGAGCTAAGAGAGCAATCCAAAGATCTGTGCAAT 181
Db 180 GCGGTGCGCTCGGTCGCGCGGAGAGCTAAGAGAGCAATCCAAAGATCTGTGCAAT 239
Qy 182 GTGTTTTGTGAGCGCGCGGGAATGTGCAGTACAGAGAGGGGAACCCACTGTCTC 241
Db 240 GTGTTTTGTGAGCGCGCGGGAATGTGCAGTACAGAGAGGGGAACCCACTGTCTC 299
```

QY 242 TGCATTGAGCAATGCACAACTCTCAAGAGGCGCTGTGTGGCGATATGCAAGACCTTAC 301  
 Db 300 TGCATTGAGCAATGCACAACTCTCAAGAGGCGCTGTGTGGCGATATGCAAGACCTTAC 359  
 QY 302 CTCAACCACTGTGAATGCATGCATGAGTGCCTGCTCATGTATCCAAATCCAGGTTAT 361  
 Db 360 CTCAACCACTGTGAATGCATGCATGAGTGCCTGCTCATGTATCCAAATCCAGGTTAT 419  
 QY 362 TACGATGACACTGCAAGAGAGAAATCCGTAAGTCCATCTGGCAAGCCAGTTGTC 421  
 Db 420 TACGATGACACTGCAAGAGAGAAATCCGTAAGTCCATCTGGCAAGCCAGTTGTC 479  
 QY 422 TATCAGTCCAACTGATGAGTGCCTGCAAGCGCATATCATGAGGCTGGAAGCTAGATC 481  
 Db 480 TATCAGTCCAACTGATGAGTGCCTGCAAGCGCATATCATGAGGCTGGAAGCTAGATC 539  
 QY 482 ATCTCAGATGCTGTTCTCTAAGGCGAGCAATCAATGAAATCTTACAGATATTT 541  
 Db 540 ATCTCAGATGCTGTTCTCTAAGGCGAGCAATCAATGAAATCTTACAGATATTT 599  
 QY 542 AAGACTTTGATATGATGATCTCGCTGAGCTCCAGTGAATTCCTGAGTTTGGA 601  
 Db 600 AAGACTTTGATATGATGATCTCGCTGAGCTCCAGTGAATTCCTGAGTTTGGA 659  
 QY 602 CAGAAATGAACTGCCATATATTTACAGATATCCAGACAGAGAGAACAAAGTTGCTT 661  
 Db 660 CAGAAATGAACTGCCATATATTTACAGATATCCAGACAGAGAGAACAAAGTTGCTT 719  
 QY 662 AGGGGACTCTGTGTTGATGCTCTCATGAACTGTCTGATGAAATTCCTGATGAAATCTC 721  
 Db 720 AGGGGACTCTGTGTTGATGCTCTCATGAACTGTCTGATGAAATTCCTGATGAAATCTC 779  
 QY 722 AGCTTCAAGAGTTCTCAAGTGCCTCAACCCATCTTCAACCTCTGAGAGAGAGTGT 781  
 Db 780 AGCTTCAAGAGTTCTCAAGTGCCTCAACCCATCTTCAACCTCTGAGAGAGAGTGT 839  
 QY 782 GCGCTGAGAGATGAACTGATGAGATGAGAGTGAAGCCGAGTGCATCTGTAACCGCTGT 841  
 Db 840 GCGCTGAGAGATGAACTGATGAGATGAGAGTGAAGCCGAGTGCATCTGTAACCGCTGT 899  
 QY 842 GTCTGTGCTGTGGAATTTGGGTCTGTACAGCCATGACCTGTGACGAGAGAGATGAG 901  
 Db 900 GTCTGTGCTGTGGAATTTGGGTCTGTACAGCCATGACCTGTGACGAGAGAGATGAG 959  
 QY 902 GGGGCGCAACCCAGACAGAGAGAGAGATGACCAATATGTCAGAGAGCTCCAAAGCAT 961  
 Db 960 GGGGCGCAACCCAGACAGAGAGAGAGATGACCAATATGTCAGAGAGCTCCAAAGCAT 1019  
 QY 962 CAGGAAACAGCTGAAAGACCAAGAGAGTGAAGCAAGAGATCTTAATGAGAGACACA 1021  
 Db 1020 CAGGAAACAGCTGAAAGACCAAGAGAGTGAAGCAAGAGATCTTAATGAGAGACACA 1079  
 QY 1022 GACCAAGTGTGATCCAGACATCTTCTCCA 1052  
 Db 1080 GACCAAGTGTGATCCAGACATCTTCTCMA 1110

RESULT 3  
 AL573353/c 1109 bp mRNA linear EST 31-MAY-2003  
 LOCUS AL573353 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS0D1042YC07 3-PRIME, mRNA sequence.  
 ACCESSION AL573353  
 VERSION AL573353.2 GI:31294707  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1109)  
 AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)  
 COMMENT On Feb 16, 2001 this sequence version replaced gi:12932513.  
 CONTACT Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 1910.r For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0D1042AB04NP1&cluster=1910.r. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0D1042AB04NP1.  
 Location/Qualifiers  
 1..1109  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1042YC07"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_1lb="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 27.3%; Score 1007.6; DB 9; Length 1109;  
 Best Local Similarity 99.1%; Pred. No. 44e-192;  
 Matches 1041; Conservative 2; Mismatches 4; Indels 3; Gaps 3;

QY 2628 TAACCCCACTCACTACCTCTTTCAGATGATGGAATTCGAAAGATTAAATCCTTGC 2687  
 Db 1089 TWAACCCCACTCACTACCTCTTTCAGATGATGGAATTCGAAAGATTAAATCCTTGC 1032  
 QY 2688 ATTAAGATCATTTGGAGAGAGATTAAGACATCAAGAGAAATTTAGAGTTGT 2747  
 Db 1031 ATTAAGATCATTTGGAGAGAGATTAAGACATCAAGAGAAATTTAGAGTTGT 972  
 QY 2748 AATGACAGGCAACTGTGAGAACTGTGTGCCAAATTAAGAAATTCCTTCACTTATCC 2807  
 Db 971 AATGACAGGCAACTGTGAGAACTGTGTGCCAAATTAAGAAATTCCTTCACTTATCC 913  
 QY 2808 TGTCTCATTTGAAGAGAGAAATTCACATTTGTTAGCATTTCAAGCTTTATGATCC 2867  
 Db 912 TGTCTCATTTGAAGAGAGAAATTCACATTTGTTAGCATTTCAAGCTTTATGATCC 853  
 QY 2868 ATCCCATTAATACTCTTCAACTCCATCTTGTCAAGTCTGAATGCAAGCTCCCTGCA 2927  
 Db 852 ATCCCATTAATACTCTTCAACTCCATCTTGTCAAGTCTGAATGCAAGCTCCCTGCA 793  
 QY 2928 AGTGCCTTGAAGAACTCAAGCAGACGCGCTTAATCAAGGTTTACAGCCCTTGAGCA 2987  
 Db 792 AGTGCCTTGAAGAACTCAAGCAGACGCGCTTAATCAAGGTTTACAGCCCTTGAGCA 733  
 QY 2988 CTATGAGAGAGGCAAGATGACCAATTTGTTAAAGCAGAGAAACCAAGTGTCTTT 3047  
 Db 732 CTATGAGAGAGGCAAGATGACCAATTTGTTAAAGCAGAGAAACCAAGTGTCTTT 673  
 QY 3048 CACTAGCATTTTGAACAATGTTTATCATCCAAAGCACTTACCTCGCAACATTTGAAC 3107  
 Db 672 CACTAGCATTTTGAACAATGTTTATCATCCAAAGCACTTACCTCGCAACATTTGAAC 613  
 QY 3108 CCAAGAGCAATCCACATCTCTTGAAGTCTGAGCTTCTGATGAAATGAGGCAAGCTGT 3167  
 Db 612 CCAAGAGCAATCCACATCTCTTGAAGTCTGAGCTTCTGATGAAATGAGGCAAGCTGT 553  
 QY 3168 CGTCTATGCGGTAGAAATCAATGATCTGAGAGACATTCATGAGAGGCTGTAATAGCTTA 3227  
 Db 552 CGTCTATGCGGTAGAAATCAATGATCTGAGAGACATTCATGAGAGGCTGTAATAGCTTA 493  
 QY 3228 GTCTGGGAGAGTCTTCATTAAGTTTTCATGAGAGCAAAACAGAGATTAACTAGGTTT 3287

492 GTCTGGGGAGCTTCCATAAGTTTTCATGGAGCAAAACAGGATTAAACATAGGTTT 433  
 3288 GGTTCCTTCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGGCTTCTTGGGCTTTCTCT 3347  
 432 GGTTCCTTCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGGCTTCTTGGGCTTTCTCT 373  
 3348 GTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTT 3407  
 372 GTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACAT 313  
 3408 CCCACATGCCGTGACTCTGGACTATATACAGTTTGTGAAAGCAGGGTTCCTCTGCCCTGCT 3467  
 312 CCCACATGCCGTGACTCTGGACTATATACAGTTTGTGAAAGCAGGGTTCCTCTGCCCTGCT 253  
 3468 AACAGCCACGTGGACAGCTGTAATGCTTCTTCCCTTACACCTATGTTTAAAGTAGTC 3527  
 252 AACAGCCACGTGGACAGCTGTAATGCTTCTTCCCTTACACCTATGTTTAAAGTAGTC 193  
 3528 AACCTTCAAGAAACAACTCTAAACAGAGTTTCTGTTGCATATGTTTGTGAACCTTGATTT 3587  
 192 AACCTTCAAGAAACAACTCTAAACAGAGTTTCTGTTGCATATGTTTGTGAACCTTGATTT 133  
 3588 GTATTAGTAGGCTTCTATATGATTAACCTTGTGTTTGTAACTCTGATTCCTCTT 3647  
 132 GTATTAGTAGGCTTCTATATGATTAACCTTGTGTTTGTAACTCTGATTCCTCTT 73  
 3648 TCGGATACTATTGATGAATAAAGAAATTA 3677  
 72 TCGGATACTATTGATGAATAAAGAAATTA 43

RESULT 4  
 BM553173 1090 bp mRNA linear EST 20-FEB-2002  
 LOCUS AGENCOURT 6572495 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5467082  
 5' mRNA sequence.  
 BM553173  
 BM553173.1 GI:18791672  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1090)  
 NIH-MGC <http://mgi.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: DCTD/DTP  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LICM1968 row: i column: 03  
 High quality sequence stop: 690.  
 Location/Qualifiers  
 1..1090  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5467082"  
 /tissue\_type="amelanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_41"

FEATURES  
 source  
 1..1090  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5467082"  
 /tissue\_type="amelanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_41"  
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; CDNA made by oligo-dr priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

ORIGIN  
 Query Match 26.8%; Score 991.2; DB 12; Length 1090;  
 Best Local Similarity 97.2%; Pred. No. 8.7e-189;  
 Matches 1049; Conservative 0; Mismatches 26; Indels 4; Gaps 4;

QY 537 ATTTAAGAACTTTGATATGTTGATTTCTCCCTGGAGTCCAGTGAATTCCTGAAGTTTG 596  
 Db 1 ATTTAAGAACTTTGATATGTTGATTTCTCCCTGGAGTCCAGTGAATTCCTGAAGTTTG 60  
 QY 597 TGGACACAAATGAACCTGCCATCAATATTACAAAGTATCCAGACGAGAGAACCAAGT 656  
 Db 61 TGGACACAAATGAACCTGCCATCAATATTACAAAGTATCCAGACGAGAGAACCAAGT 120  
 QY 657 TGCTTAGGGAGCTCTGTGTTGATGCTCTCATTTGAATGTTCTGTATGAAATGCTGATTGGA 716  
 Db 121 TGCTTAGGGAGCTCTGTGTTGATGCTCTCATTTGAATGTTCTGTATGAAATGCTGATTGGA 180  
 QY 717 AACTCAGCTTCCAAAGAGTTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCTCGAGAAGA 776  
 Db 181 AACTCAGCTTCCAAAGAGTTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCTCGAGAAGA 240  
 QY 777 AGTGTCCCTCGAAGATGAACAGTATGAGATGAGCTGAGACCGAGGTGACTGTAAACC 836  
 Db 241 AGTGTCCCTCGAAGATGAACAGTATGAGATGAGCTGAGACCGAGGTGACTGTAAACC 300  
 QY 837 GCTGTGCTGTGCTGTGGAATTTGGGTCTGTACAGCATGACCTGTGACCGAAGAAATC 896  
 Db 301 GCTGTGCTGTGCTGTGGAATTTGGGTCTGTACAGCATGACCTGTGACCGAAGAAATC 360  
 QY 897 AGAAGGGGGCCAGACCCAGACAGAGGAGAGATGACAGATATGTCAGAGGTCTCCAAA 956  
 Db 361 AGAAGGGGGCCAGACCCAGACAGAGGAGAGATGACAGATATGTCAGAGGTCTCCAAA 420  
 QY 957 AGCATCAGAAACAGCTGAAAGCCAGAGAGTGCACCAACCAAGAGATCTAATGAGGAG 1016  
 Db 421 AGCATCAGAAACAGCTGAAAGCCAGAGAGTGCACCAACCAAGAGATCTAATGAGGAG 480  
 QY 1017 GCACAGACAGTGTCTGATCCAGCATCTTCTCCACTTTCACGCTGAGTTCAGTATACA 1076  
 Db 481 GCACAGACAGTGTCTGATCCAGCATCTTCTCCACTTTCAGCTGAGTTCAGTATACA 540  
 QY 1077 CAAAGTGTCTGATAGTGCCTGCAATATCAGCATATTTGCTTATATAGCATGATTTATT 1136  
 Db 541 CAAAGTGTCTGATAGTGCCTGCAATATCAGCATATTTGCTTATATAGCATGATTTATT 600  
 QY 1137 TTGTTTATTTGTTTTCATTAAGGATATGAAGTGGCTGGCTAGGAGGAGGAGGCGCAC 1196  
 Db 601 TTGTTTATTTGTTTTCATTAAGGATATGAAGTGGCTGGCTAGGAGGAGGAGGCGCAC 660  
 QY 1197 AGCCTTCATTTCTAGAGTGTCTTAAAGAAACTGTAAATGGTGTCTCTGGGGCTGGAGGC 1256  
 Db 661 AGCCTTCATTTCTAGAGTGTCTTAAAGAAACTGTAAATGGTGTCTCTGGGGCTGGAGGC 720  
 QY 1257 TAGTAAGGAAACTGCATCAGCATGAAAGAGGAAACAGACCCCAATCTGAACCTCTTTGA 1316  
 Db 721 TAGTAAGGAAACTGCATCAGCATGAAAGAGGAAACAGACCCCAATCTGAACCTCTTTGA 780  
 QY 1317 GTTTACTCATCTGTACAGAGCTGCAGGGAGTGCACAGATGCCAGAGAGAACTTAGCA 1376  
 Db 781 GTTTACTCATCTGTACAGAGCTGCAGGGAGTGCACAGATGCCAGAGAGAACTTAGCA 840  
 QY 1377 GGGTCTCCCCGGAG-GAGAGGTTTGGGAAGCTCCACGGAGAGGAAACGCTCTCTCTCCA 1435  
 Db 841 GGGTCTCCCCGGAGAGAGTTTGGGAGGCTCCACGAGAGGAAACGCTCTCTCTCCA 900  
 QY 1436 GCGCTTTTCCATTGCCGTGACATGACAGACTCCAGCATCCAGCATCTCTTGGTCCCA 1495  
 Db 901 GCGCTTTTCCATTGCCGTGACATGACAGACTCCAGCATCCAGCATCAAGGCATCTCTTGGTCCCA 960

QY 1496 ATAACTGCTCTAGATACATAGCCATACCTAGTCTAACCAGTGTCCCTCAGACTT-GG 1554  
 DB 961 ATAACTGCTCTCTA-AAACATAGCCATCTGCTAGNTAAACCAGGTCCTTCAAACTTGGG 1019  
 QY 1555 ATGAGTCTTCT-GGAGGGTACACCCAAATGATCAGATACCTTGTATACATTTTGAGGCC 1612  
 DB 1020 ATGAGTCTTCTGGGAGGGAACCCAAATGATGCGCATCTTGAATAATTTGAGGCC 1078

RESULT 5  
 BM463192 1094 bp mRNA linear EST 05-FEB-2002  
 LOCUS AGENCOURT 6447506 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5537917  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM463192  
 VERSION BM463192.1 GI:18512234  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1094)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC/DCTD/DTF  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 plate: LLAM12229 row: p column: 14  
 High quality sequence stop: 675.

FEATURES  
 source  
 1..1094  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5537917"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_72"  
 /note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 2 kb. Library constructed by Life  
 Technologies."

ORIGIN  
 Query Match 26.8%; Score 988.6; DB 12; Length 1094;  
 Best Local Similarity 96.0%; Pred. No. 2.9e-189;  
 Matches 1031; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

QY 2611 TCGACTTCTGTGATTAATTAACCCCACTCAGTACCTGTTTCAGATGCAATGGAAATACCA 2670  
 DB 1 TCGACTTCTGTGATTAATTAACCCCACTCAGTACCTGTTTCAGATGCAATGGAAATACCA 60  
 QY 2671 AAGATTAATCTTGACATAGATCTCATTTGCGAAGAGCAGATTAAGACCATCAGAAG 2730  
 DB 61 AAGATTAAGCCCTTGACATAGATCTCATTTGCGAAGAGCAGATTAAGACCATCAGAAG 120  
 QY 2731 GAAATTAATTTAGTGTGTAATGACAGGCAACTGTGAGAAACTGTGTGCCAAAATAGAA 2790  
 DB 121 GAAATTAATTTAGTGTGTAATGACAGGCAACTGTGAGAAACTGTGTGCCAAAATAGAA 180  
 QY 2791 TTCCTCTAGTTTTCTTGTCTCATTTGAAGAGAGAAATTCCTTGTGTAGCATTT 2850  
 DB 181 TTCCTCTAGTTTTCTTGTCTCATTTGAAGAGAGAAATTCCTTGTGTAGCATTT 240  
 QY 2851 CAAGCTTTTATGTATCCATCCCACTTAAAACTTCCAACTCCCACTTGTTCAGTCTGAA 2910

DB 241 CBAAGCTTTATGTATCCATCCATCTAAAAACTCTTCAAACTCCACTTGTTCAGTCTGAA 300  
 QY 2911 ATGCACTCCCTGTCCTGCAAGTGCCTTGGAGAACTCAACAGCAGACGCTTAACTCAAGGTT 2970  
 DB 301 ATGCACTCCCTGTCCTGCAAGTGCCTTGGAGAACTCAACAGCAGACGCTTAACTCAAGGTT 360  
 QY 2971 TTACAGAGCCCTTGGACACTATGGGAGGAGGCAAGAGTACACCAATTTGTTAAAGCAAG 3030  
 DB 361 TTACAGAGCCCTTGGACACTATGGGAGGAGGCAAGAGTACACCAATTTGTTAAAGCAAG 420  
 QY 3031 AAACACAGTGTCTTCTTCACTAGTCAATTTAGAAACATGGTATATCATCCAGACTACTCTAC 3090  
 DB 421 AAACACAGTGTCTTCTTCACTAGTCAATTTAGAAACATGGTATATCATCCAGACTACTCTAC 480  
 QY 3091 COTGCAACATTTGAATCCCAAGAGCAAAATCCACATTTCTCTTGTAGTTCTGAGGCTTCTGT 3150  
 DB 481 COTGCAACATTTGAATCCCAAGAGCAAAATCCACATTTCTCTTGTAGTTCTGAGGCTTCTGT 540  
 QY 3151 GTPAATAGGCGAGTGTCTGTATGCCGTAGAAATCAATGATCTGAGGACCATTCATGGA 3210  
 DB 541 GTPAATAGGCGAGTGTCTGTATGCCGTAGAAATCAATGATCTGAGGACCATTCATGGA 600  
 QY 3211 AGCTGCTAAATAGCCCTAGTCTGGGAGGCTTCCATAAAAGTTTTCATCGAGCAGCAAAAC 3270  
 DB 601 AGCTGCTAAATAGCCCTAGTCTGGGAGGCTTCCATAAAAGTTTTCATCGAGCAGCAAAAC 660  
 QY 3271 AGGATTAAACTAGTGTGTTTCTTTCAGCCCTCTAAAGCATAGGGCTTAGCTGCAGGC 3330  
 DB 661 AGGATTAAACTAGTGTGTTTCTTTCAGCCCTCTAAAGCATAGGGCTTAGCTGCAGGC 720  
 QY 3331 TTCCCTGGCTTCTCTGCTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGATCCA 3390  
 DB 721 TTCCCTGGCTTCTCTGCTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGATCCA 780  
 QY 3391 GTGTCCATGGAACCTTCCACATGCCGTGACTCTGGAATATATCAGTTTGTGGAAAGA 3450  
 DB 781 GTGTCCATGGAACCTTCCACATGCCGTGACTCTGGAATATATCAGTTTGTGGAAAGA 840  
 QY 3451 GGGTTCCTGCTGCTGTAACAGCCCACTGAGCAGCTGTAATCTTTCCTTTACACC 3510  
 DB 841 GGGTTCCTGCTGCTGTAACAGCCCACTGAGCAGCTGTAATCTTTCCTTTACACC 900  
 QY 3511 TATGTTTTTAAGTAGTCAAACTTCAAGAAACAATCTAAACAGTTTCTGTTGCATATGTG 3570  
 DB 901 TATGTTTTTAAGTAGTCAAACTTCTCAG-AACCATCTAAACAGTTTCTGTTGCATATGTG 959  
 QY 3571 TTGTGAATCTGATTTGATTTAGTAGCTCTTATATGCA-TTAACTGTTTGTGTA 3629  
 DB 960 TTGTGAATCTGATTTGATTTAGTAGCTCTTATATGCA-TTAACTGTTTGTGTA 1019  
 QY 3630 ACTCCTGATTTCTTCTTTTCGGATACTATTGATGAATAAGAAATTAAGTGAA 3683  
 DB 1020 ACTCCTGATTTCTTCTTTTCGGATACTATTGATGAATAAGAAATTAAGTGAA 1073

RESULT 6  
 BC006185 1677 bp mRNA linear HTC 17-DEC-2003  
 LOCUS Mus musculus follistatin-like, mRNA (cdna clone IMAGE:3498052).  
 DEFINITION  
 ACCESSION BC006185  
 VERSION BC006185.1 GI:13544115  
 KEYWORDS HTC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1677)  
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,







```
2y 1069 ACTCGCCATAGGCAATTCATCAGGATTAATGGTTTCCCGTTCTGAAATTCACATGAATGAT 1128
Db 811 ACAGATGATCATACCCCTCATCAGAGTCATGGTTTCCAGGAGTGAGATTGATCTGTTTAAC 870
2y 1129 ATCAAAGCAATTCATCAGAGATGATGGTATCTCCCTTTGCCAAGCATTCTCGGATGAA 1188
Db 871 ATCAGGAAGGATTTAGGAAGATTTGCCACCTCTCTTTATCCATGATTAAGGGAT 930
2y 1189 ACCAAAGAGATGATGAAGAAATCCTGGTGGCTCTTTGTGGAGAAA 1235
Db 931 ACATCTGGGACTATAAGAAAGCTCTTCTGTCTCTCCGGAGAAGA 977

RESULT 13
US-09-324-096A-3
; Sequence 3, Application US/09324096A
; Patent No. 6323313
; GENERAL INFORMATION:
; APPLICANT: Tait, Jonathan
; TITLE OF INVENTION: ANNEXIN DERIVATIVE WITH ENDOGENOUS CHELATION SITES
; FILE REFERENCE: UOEW-1-13841
; CURRENT APPLICATION NUMBER: US/09/324,096A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(981)
US-09-324-096A-3

Query Match 15.6%; Score 236.6; DB 4; Length 981;
Best Local Similarity 53.1%; Pred. No. 2.5e-62;
Matches 503; Conservative 0; Mismatches 444; Indels 0; Gaps 0;

2y 289 GGTCCCGATCAGCGGTGAGCCCTCTACCTTCAATCATCATCTCGGATGCGTGCC 348
Db 31 GTTCTCAGAGCACTGTGACTGACTTCCCTGGATTGATGAGCGGGCTGATGAGAACT 90
2y 349 TTGCATAGGCATTAATGGTTAAAGGTGTGATGAAGCAACCATTCATTCATTAAT 408
Db 91 CTTCGGAAGGCTATGAAGGCTTGGGCACAGATGAGGAGCATCTCTGACCTGTTGACA 150
2y 409 AAGCGAACAATGCACAGCGTCAACAGATCAAAAGCAATATCTCCAGGAAACAGGAAAG 468
Db 151 TCCCGAAGTAAATGCTCAGCGCCAGGAAATCTCTGAGCTTTTAAAGCTGTTGGCAGG 210
2y 469 CCCCTGGATGAAAGCTTTCGGAACGCTTTGCTTCTCTGTAAGGCTGACCGGCTCTGAG 768
Db 451 GACACTTCAGGCTACTACCGGAGTGTGTGGTCTCTCTCTCCTTCAGGCTAACAGACCT 510
2y 769 GACTTTGGTGTGAATGAAGAATTGGCTGANTTCAGATGCCGGGCTTGTATGAAGCAGGA 828
```

```
Db 511 GATGCTGGAATTCATGAAGCTCAAGTTGAAACAAGATGCTCAGGCTTTATTTTCAGGCTGA 570
2y 829 GAAAGGAGAAAGGGGACAGAGCTTAAACGTGTTCATACCATCTCTTACCCAGAGACTAT 888
Db 571 GAACTTAAATGGGGACAGATGAAGAAAGTTTATCACCATCTTTGGAAACACGAAAGTGTG 630
2y 889 CCACAACCTTCGAGAGTCTTTTCAGAAATACACCAAGTACAGTAAAGCATGACATGAACAAA 948
Db 631 TCTCATTTGAGAAAGGTGTTTGACAAGTACATGACTATATCAGGATTTCAAATTTGAGGAA 690
2y 949 GTTCTGGACCTGGAGTTGAAAGGTGACATTTGAGAAATCCCTCACAGCTATCTGTGAAGTGC 1008
Db 691 ACCATTGACCGGAGACTTCTGGCAATTTAGAGCAACTACTCTCTGCTGTTGTAATCT 750
2y 1009 GCCACAAGCAACCAAGCTTCTTTTCAGAGAGCTTTCATCAAGCCATGAAGGCTTTGGA 1068
Db 751 ATTGGAAGTATACCTGCTTACCTTTCGAGAGACCTCTTATATGCTATGAAGGAGCTGGG 810
2y 1069 ACTCGCCATAGGCAATTCATCAGGATTAATGGTTTCCCGTTCTGAAATTCACATGAATGAT 1128
Db 811 ACAGATGATCATACCCCTCATCAGAGTCATGGTTTCCAGAGTGAGATTGATCTGTTTAAC 870
2y 1129 ATCAAAGCAATTCATCAGAGATGATGGTATCTCCCTTTGCCAAGCCATCTCTGATGAA 1189
Db 871 ATCAGGAAGGATTTAGGAAGAAATTTGCCACCTCTCTTTATTCATGATTAAGGGAT 930
2y 1189 ACCAAAGAGAGATGATGAAGAAATCCTGGTGGCTCTTTGTGGAGAAA 1235
Db 931 ACATCTGGGACTATAAGAAAGCTCTTCTGTGCTCTCCGGAGAAGA 977

RESULT 14
US-09-324-096A-5
; Sequence 5, Application US/09324096A
; Patent No. 6323313
; GENERAL INFORMATION:
; APPLICANT: Tait, Jonathan
; TITLE OF INVENTION: ANNEXIN DERIVATIVE WITH ENDOGENOUS CHELATION SITES
; FILE REFERENCE: UOEW-1-13841
; CURRENT APPLICATION NUMBER: US/09/324,096A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(981)
US-09-324-096A-5

Query Match 15.6%; Score 236.6; DB 4; Length 981;
Best Local Similarity 53.1%; Pred. No. 2.5e-62;
Matches 503; Conservative 0; Mismatches 444; Indels 0; Gaps 0;

2y 289 GGTCCCGATCAGCGGTGAGCCCTCTACCTTCAATTCATTCCTCGATGCTGCTGCC 348
Db 31 GTTCTCAGAGCACTGTGACTGACTTCCCTGGATTGATGAGCGGGCTGATGAGAACT 90
2y 349 TTGCATAGGCATTAATGGTTAAAGGTGTGATGAAGCAACCATTCATTCATTAAT 408
Db 91 CTTCGGAAGGCTATGAAGGCTTGGGCACAGATGAGGAGCATCTCTGACCTGTTGACA 150
2y 409 AAGCGAACAATGCACAGCGTCAACAGATCAAAAGCAATATCTCCAGGAAACAGGAAAG 468
Db 151 TCCCGAAGTAAATGCTCAGCGCCAGGAAATCTCTGAGCTTTTAAAGCTGTTGGCAGG 210
2y 469 CCCCTGGATGAAAGCTTTCGGAACGCTTTGCTTCTCTGTAAGGCTGACCGGCTCTGAG 768
Db 451 GACACTTCAGGCTACTACCGGAGTGTGTGGTCTCTCTCTCCTTCAGGCTAACAGACCT 510
2y 769 GACTTTGGTGTGAATGAAGAATTGGCTGANTTCAGATGCCGGGCTTGTATGAAGCAGGA 828
```



GenCore version 5.1.6  
Copyright (C) 1993 - 2004 Compugen Ltd.

2M nucleic - nucleic search, using sw model

Run on: March 11, 2004, 03:37:21 ; Search time 353.257 Seconds  
(without alignments)  
15799.545 Million cell updates/sec

Title: US-10-084-817-8

Perfect score: 1516

Sequence: 1 cttgtttttggacatagct.....gctgaataataatgncgrc 1516

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgm2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgm2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgm2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgm2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgm2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgm2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgm2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgm2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgm2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgm2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgm2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgm2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgm2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgm2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgm2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgm2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgm2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgm2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1515	99.9	1516	9	US-09-919-172-38
2	1515	99.9	1516	9	US-09-974-288-92
3	1515	99.9	1516	14	US-10-084-817-8
4	1389.4	91.6	1399	9	US-09-864-864-321
5	1387.8	91.5	1399	9	US-09-954-456-2119
6	1387.8	91.5	1399	14	US-10-240-965-2
7	1387.8	91.5	1399	14	US-10-269-909-2
8	1387.8	91.5	1399	15	US-10-133-937-2
9	1387.8	91.5	1399	15	US-10-159-563-2
10	1379.4	91.0	1431	9	US-09-925-301-46
11	1367.4	90.2	1377	14	US-10-210-120-18
12	886.8	58.5	1402	9	US-09-917-800A-1585
13	886.8	58.5	1402	14	US-10-316-253-299
14	822	54.2	2500	10	US-09-814-353-20447
15	597.4	39.4	600	12	US-10-085-783A-51438

16	597.4	39.4	600	15	US-10-242-535A-51438
17	589.4	38.9	614	9	US-09-864-864-321
18	583.4	38.5	618	12	US-10-085-783A-13898
19	583.4	38.5	618	15	US-10-242-535A-13898
20	529	34.9	592	12	US-10-085-783A-48352
21	529	34.9	592	15	US-10-242-535A-48352
22	519.4	34.3	537	9	US-09-777-584-134
23	519.4	34.3	537	14	US-10-015-219-134
24	505	33.3	557	12	US-10-085-783A-48313
25	505	33.3	557	15	US-10-242-535A-48313
26	500	33.0	530	12	US-10-085-783A-54108
27	500	33.0	530	15	US-10-242-535A-54108
28	496	32.7	590	12	US-10-085-783A-48227
29	496	32.7	590	15	US-10-242-535A-48227
30	491.4	32.4	500	12	US-10-085-783A-23999
31	491.4	32.4	500	15	US-10-242-535A-23999
32	482	31.8	495	12	US-10-085-783A-38267
33	482	31.8	495	15	US-10-242-535A-38267
34	474	31.3	499	12	US-10-085-783A-26128
35	474	31.3	499	15	US-10-242-535A-26128
36	468.2	30.9	473	12	US-10-085-783A-24370
37	468.2	30.9	473	15	US-10-242-535A-24370
38	460.2	30.4	470	12	US-10-085-783A-42942
39	460.2	30.4	470	15	US-10-242-535A-42942
40	459.4	30.3	461	9	US-09-998-598-607
41	458	30.2	473	12	US-10-085-783A-31104
42	458	30.2	473	15	US-10-242-535A-31104
43	457.8	30.2	461	9	US-09-998-598-1129
44	455.2	30.0	472	12	US-10-085-783A-55749
45	455.2	30.0	472	15	US-10-242-535A-55749

ALIGNMENTS

RESULT 1

US-09-919-172-38  
; Sequence 38, Application US/09919172  
; Patent No. US20020119463A1  
; GENERAL INFORMATION:  
; APPLICANT: Paris, Mary  
; APPLICANT: Turner, Christopher M.  
; TITLE OF INVENTION: PROSTATE CANCER MARKERS  
; FILE REFERENCE: PA-0036 US  
; CURRENT APPLICATION NUMBER: US/09/919,172  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/222,469  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PERL Program  
; SEQ ID NO 38  
; LENGTH: 1516  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020119463A1 1303785CB1  
; NAME/KEY: unsure  
; LOCATION: 1512  
; OTHER INFORMATION: a, t, c, g, or other

US-09-919-172-38

Query Match 99.9%; Score 1515; DB 9; Length 1516;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CTTTGTTTTGGACATAGCTGAGCCCTTACTTCAACAGAGGCGCAATTAACCT 60
'Db	1	CTTTGTTTTGGACATAGCTGAGCCCTTACTTCAACAGAGGCGCAATTAACCT 60
QY	61	TCGGTGTCTAGCTGGCTTCTTTAAATCTATAAATCAGAGCCCACTCTCCAC 120
Db	61	TCGGTGTCTAGCTGGCTTCTTTAAATCTATAAATCAGAGCCCACTCTCCAC 120

121 TGCCAGTGTGAATCTTTCAGAGAGAAATTCCTTTAGTTCTTTTCAGAGAGGTAGAGAT 180  
121 TGCCAGTGTGAATCTTTCAGAGAGAAATTCCTTTAGTTCTTTTCAGAGAGGTAGAGAT 180  
181 AAAGACACTTTTCAAAAATGGCAATGGTATCAGAAATCCTCAAGCAGGCCCTGGTTTATT 240  
181 AAAGACACTTTTCAAAAATGGCAATGGTATCAGAAATCCTCAAGCAGGCCCTGGTTTATT 240  
241 GAAATGAAGAGCAGCAATATGTTCAAACTGTGAGTCAATCCAAAGGTGGTCCCGGATCA 300  
241 GAAATGAAGAGCAGCAATATGTTCAAACTGTGAGTCAATCCAAAGGTGGTCCCGGATCA 300  
301 GCGGTGAGCCCTTATCCTACCTTCAATCCATCCTCGATGTGCTGCTTGCCTTGCATGAAGGCC 360  
301 GCGGTGAGCCCTTATCCTACCTTCAATCCATCCTCGATGTGCTGCTTGCCTTGCATGAAGGCC 360  
361 ATAAATGAAGGTGTGATGAAGCAACCAATGATGACATTTCAATTAAGGGAACCAAT 420  
361 ATAAATGAAGGTGTGATGAAGCAACCAATGATGACATTTCAATTAAGGGAACCAAT 420  
421 GCACAGCGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGAAAGCCCTGGATGAA 480  
421 GCACAGCGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGAAAGCCCTGGATGAA 480  
481 ACACTGAAGAAAGCCCTTACAGTCACTTACCTGAGGAGTTGTTTACGCTGCTGCTTAAACT 540  
481 ACACTGAAGAAAGCCCTTACAGTCACTTACCTGAGGAGTTGTTTACGCTGCTGCTTAAACT 540  
541 CCAGCGCAATTTGATGCTGATGAATCTCGTGCCTGATGAAGGCCCTTGGAACTGATGAA 600  
541 CCAGCGCAATTTGATGCTGATGAATCTCGTGCCTGATGAAGGCCCTTGGAACTGATGAA 600  
601 GATACCTAAATGAGATTTTGGATCAAGAACTAACAAGAAATCAGAGACATTAACAGG 660  
601 GATACCTAAATGAGATTTTGGATCAAGAACTAACAAGAAATCAGAGACATTAACAGG 660  
661 GTCTACAGAGAGAACTGAAGAGAGATCTGCGCAAGACATACCTCAGACACATCTGGA 720  
661 GTCTACAGAGAGAACTGAAGAGAGATCTGCGCAAGACATACCTCAGACACATCTGGA 720  
721 GATTTGCGAAGCTTTGCTTTCTTTGCTAAGGTTGACCGATCTGAGACATTTGCTGTTG 780  
721 GATTTGCGAAGCTTTGCTTTCTTTGCTAAGGTTGACCGATCTGAGACATTTGCTGTTG 780  
781 AATGAGACTTTGCTGATTCAGATGCGAGGCTTGTATGAGCAGGAGAAAGGAGAAAG 840  
781 AATGAGACTTTGCTGATTCAGATGCGAGGCTTGTATGAGCAGGAGAAAGGAGAAAG 840  
841 GGGACAGCTAAACGTTTCAATACCATCTTACCAAGAGCTATCCACAACTTCGC 900  
841 GGGACAGCTAAACGTTTCAATACCATCTTACCAAGAGCTATCCACAACTTCGC 900  
901 AGAGTGTTCAGAAATACCAAGTACAGTCAAGCAGTCAATGAAGTTCTTGACCTG 960  
901 AGAGTGTTCAGAAATACCAAGTACAGTCAAGCAGTCAATGAAGTTCTTGACCTG 960  
961 GAGTTGAAAGGTGACATTTGAGAAATGCTTCAAGCTATCGTGAAGTGGCGCAACAA 1020  
961 GAGTTGAAAGGTGACATTTGAGAAATGCTTCAAGCTATCGTGAAGTGGCGCAACAA 1020  
1021 CCAGCTTTCTTTCAGAGAGAGCTTCAAGCCATGAAGGTTTGGAACTCGCCATAG 1080  
1021 CCAGCTTTCTTTCAGAGAGAGCTTCAAGCCATGAAGGTTTGGAACTCGCCATAG 1080  
1081 GCATTTGAGGATTAAGTTTCCCGTTCTGAAATTTGACATGAATGATATCAAGCATTC 1140  
1081 GCATTTGAGGATTAAGTTTCCCGTTCTGAAATTTGACATGAATGATATCAAGCATTC 1140  
1141 TATCAGAGATGATGATGATCTCTTTCAGAGCCATCTCGATGAAGCAAGAGAGAG 1200  
1141 TATCAGAGATGATGATGATCTCTTTCAGAGCCATCTCGATGAAGCAAGAGAGAG 1200

1201 TATGAGAAATCCTGTGGCTCTTTTGGAGGAACTAAACATTCCTTTGATGGTCTCAA 1260  
1201 TATGAGAAATCCTGTGGCTCTTTTGGAGGAACTAAACATTCCTTTGATGGTCTCAA 1260  
1261 GCTATGATCAGAGACTTTAAATTTATATTTTTCATCTTAAAGCTTAAATAGGAAAGTTT 1320  
1261 GCTATGATCAGAGACTTTAAATTTATATTTTTCATCTTAAAGCTTAAATAGGAAAGTTT 1320  
1321 CTTCAACAGCAATTTACAGTGTAGCTACATCTCTGAAATAATAGCCTTTAAATCATTT 1380  
1321 CTTCAACAGCAATTTACAGTGTAGCTACATCTCTGAAATAATAGCCTTTAAATCATTT 1380  
1381 TTATTTTAACTCTGTATTAATAGAGATAAGTCCATTTTAAATAGTTTCCCAAC 1440  
1381 TTATTTTAACTCTGTATTAATAGAGATAAGTCCATTTTAAATAGTTTCCCAAC 1440  
1441 CATAAAAACCTTATACAGTTGTTCTAGTAAACAATACATCAGAAAGATGTCTATGTAGCTG 1500  
1441 CATAAAAACCTTATACAGTTGTTCTAGTAAACAATACATCAGAAAGATGTCTATGTAGCTG 1500  
1501 AAAATAAAATGNCGTC 1516  
1501 AAAATAAAATGNCGTC 1516

RESULT 2  
US-09-974-298-92  
; Sequence 92, Application US/09974298  
; Patent No. US20020156263A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Huei-Mei  
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER  
; FILE REFERENCE: PA-0037 P  
; CURRENT APPLICATION NUMBER: US/09/974,298  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: 60/238,331  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 194  
; SOFTWARE: PERL Program  
; SEQ ID NO 92  
; LENGTH: 1516  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1303785CB1  
; NAME/KEY: unsure  
; LOCATION: 1512  
; OTHER INFORMATION: a, t, c, g, or other  
; US-09-974-298-92

Query Match 99.9%; Score 1515; DB 9; Length 1516;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTTTGTGTTTGGACATAGCTGAGCCATGTACTTCAACAGAGCGCCCAATTAACCT 60  
1 CTTTGTGTTTGGACATAGCTGAGCCATGTACTTCAACAGAGCGCCCAATTAACCT 60  
61 TCTGTTGCTAGGTGTGGCTTCTTTTAAATTCCTTATATAATCAGAGCCCAAGTCTCCAC 120  
61 TCTGTTGCTAGGTGTGGCTTCTTTTAAATTCCTTATATAATCAGAGCCCAAGTCTCCAC 120  
121 TCCAGTGTGAATCTTCAGAGAGAAATTTCTTTTGTAGTTCTTTTGCAGAGGTAGAGAT 180  
121 TCCAGTGTGAATCTTCAGAGAGAAATTTCTTTTGTAGTTCTTTTGCAGAGGTAGAGAT 180  
181 AAAGACACTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCCCTGGTTTATT 240  
181 AAAGACACTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCCCTGGTTTATT 240  
241 GAAATGAAGAGCAGCAATATGTTCAAACTGTGAGTCAATCCAAAGGTGGTCCCGGATCA 300



241 GAAATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCATCCAAAGGTGGTCCCGATCA 300  
 301 GCGGTGAGCCCTTATCTCTACCTTCAATCCATCTCGATGTCGCTCTTGCATAGGCC 360  
 301 GCGGTGAGCCCTTATCTCTACCTTCAATCCATCTCGATGTCGCTCTTGCATAGGCC 360  
 361 ATAATGGTTAAAGGTGTGGATGAAGCAACCATCATTTGACATTTAACTAAGCGAAACAT 420  
 361 ATAATGGTTAAAGGTGTGGATGAAGCAACCATCATTTGACATTTAACTAAGCGAAACAT 420  
 421 GCACAGGCTCAACAGATCAAGCAGATATCTCCAGGAAACAGGAAAGCCCTGGATGAA 480  
 421 GCACAGGCTCAACAGATCAAGCAGATATCTCCAGGAAACAGGAAAGCCCTGGATGAA 480  
 481 ACACGAGAAAGCCCTTACAGGTCACCTTGAAGAGGTGTTTGTAGCTCTGTCTAAACT 540  
 481 ACACGAGAAAGCCCTTACAGGTCACCTTGAAGAGGTGTTTGTAGCTCTGTCTAAACT 540  
 541 CCAGCGCAATTTGATGCTGATGAATCTCGCTGCTGCCATGAAGGCCCTTGGACTGATGA 600  
 541 CCAGCGCAATTTGATGCTGATGAATCTCGCTGCTGCCATGAAGGCCCTTGGACTGATGA 600  
 601 GATACCTCTAATTTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGACATTAACAGG 660  
 601 GATACCTCTAATTTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGACATTAACAGG 660  
 661 GTCTACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATTAACCTCAGACATCTGGA 720  
 661 GTCTACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATTAACCTCAGACATCTGGA 720  
 721 GATTTTCGGAACGCTTTGCTTTCTTCTGCTAAGGAGTACCGCATCTGAGGACTTTGGTGTG 780  
 721 GATTTTCGGAACGCTTTGCTTTCTTCTGCTAAGGAGTACCGCATCTGAGGACTTTGGTGTG 780  
 781 AATGAAGACTTTGGTGTGATTCAGATGCGCAGGCGCTTGTATGAAGCAGAGAAAGGAGAG 840  
 781 AATGAAGACTTTGGTGTGATTCAGATGCGCAGGCGCTTGTATGAAGCAGAGAAAGGAGAG 840  
 841 GGGACAGAGCTTAAAGCTGTTTCAATACCATCTCTACCAAGAGAGCTATCACAACCTTCCG 900  
 841 GGGACAGAGCTTAAAGCTGTTTCAATACCATCTCTACCAAGAGAGCTATCACAACCTTCCG 900  
 901 AGAGTGTTCAGAAATACCAAGTACAGTAAGTAAGCATGACATGAACAAAGTCTGGACCTG 960  
 901 AGAGTGTTCAGAAATACCAAGTACAGTAAGTAAGCATGACATGAACAAAGTCTGGACCTG 960  
 961 GAGTTGAAGGTGACATTTGAGAAATGCTTCAAGCTATCGTGAAGTGGCGCAACAGCAA 1020  
 961 GAGTTGAAGGTGACATTTGAGAAATGCTTCAAGCTATCGTGAAGTGGCGCAACAGCAA 1020  
 1021 CCAGCTTTCTTTGACAGAGAGCTTCAACAGCCATGAAGGTGTTGAACTCGCCATAAG 1080  
 1021 CCAGCTTTCTTTGACAGAGAGCTTCAACAGCCATGAAGGTGTTGAACTCGCCATAAG 1080  
 1081 GCATTGATCAGGATTTATGGTTTCCGTTCTGAATTCGATGATGATATCAAGAGCATTC 1140  
 1081 GCATTGATCAGGATTTATGGTTTCCGTTCTGAATTCGATGATGATATCAAGAGCATTC 1140  
 1141 TATCAGAAAGATGATGTTGATCTCCCTTTGCCAAGCCATCTCGGATGAAACCAAGAGGAG 1200  
 1141 TATCAGAAAGATGATGTTGATCTCCCTTTGCCAAGCCATCTCGGATGAAACCAAGAGGAG 1200  
 1201 TATGAGAAATCTCGTGGTCTTTGTTGGAGAACTTAAACATTTCCCTTGTGATGCTCAA 1260  
 1201 TATGAGAAATCTCGTGGTCTTTGTTGGAGAACTTAAACATTTCCCTTGTGATGCTCAA 1260  
 1261 GCTATGATCAGAAAGCTTTAATATATATTTTCACTTATAGCTTAAATAGGAAGTTT 1320  
 1261 GCTATGATCAGAAAGCTTTAATATATATTTTCACTTATAGCTTAAATAGGAAGTTT 1320  
 1321 CTTCAACAGGATTAAGTGTAGCTACCTACATGCTGGAATAATATAGCCCTTTAAATCATTT 1380  
 1321 CTTCAACAGGATTAAGTGTAGCTACCTACATGCTGGAATAATATAGCCCTTTAAATCATTT 1380

QY 1381 TTATATTATAACTCTGTATATAAGATAAGTCAATTTTAAAAATGTTTTCCTCCAAAC 1440  
 Db 1381 TTATATTATAACTCTGTATATAAGATAAGTCAATTTTAAAAATGTTTTCCTCCAAAC 1440  
 QY 1441 CATAAACCCCTATACAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTTCTATAGCTG 1500  
 Db 1441 CATAAACCCCTATACAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTTCTATAGCTG 1500  
 QY 1501 AAAATAAAATGNCCTC 1516  
 Db 1501 AAAATAAAATGNCCTC 1516

RESULT 3  
 US-10-084-817-8  
 ; Sequence 8, Application US/10084817  
 ; Publication No. US20030119009A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Susan Stuart  
 ; APPLICANT: Jed G. Nuchtern  
 ; APPLICANT: Sharon E. Plon  
 ; APPLICANT: Jason M. Shohet  
 ; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION  
 ; FILE REFERENCE: PA-0046 US  
 ; CURRENT APPLICATION NUMBER: US/10/084,817  
 ; PRIOR FILING DATE: 2002-02-25  
 ; PRIOR FILING DATE: 2001-02-23  
 ; NUMBER OF SEQ ID NOS: 365  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 8  
 ; LENGTH: 1516  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No. US20030119009A1 1303785CB1  
 ; NAME/KEY: unsure  
 ; LOCATION: 1512  
 ; OTHER INFORMATION: a, t, c, g, or other  
 US-10-084-817-8

Query Match 99.9%; Score 1515; DB 14; Length 1516;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTTTGTGTTTGGACATAGCTGAGCCATGTACTTCAACAGAGGCGAGCCAAATTACTAACT 60  
 Db 1 CTTTGTGTTTGGACATAGCTGAGCCATGTACTTCAACAGAGGCGAGCCAAATTACTAACT 60  
 QY 61 TCTGGTTGCTAGGTGCTGCTTCTTAAATCCCTATAAATCAGAGCCAGCTCTCCAC 120  
 Db 61 TCTGGTTGCTAGGTGCTTCTTAAATCCCTATAAATCAGAGCCAGCTCTCCAC 120  
 QY 121 TGCCAGTGTGAATCTTCAAGAGAAATTTCTTCTTAGTTCCTTCAAGAGGTAGAGAT 180  
 Db 121 TGCCAGTGTGAATCTTCAAGAGAAATTTCTTCTTAGTTCCTTCAAGAGGTAGAGAT 180  
 QY 181 AAGACACATTTTCAAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGCCCTGGTTATT 240  
 Db 181 AAGACACATTTTCAAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGCCCTGGTTATT 240  
 QY 241 GAAATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGATCA 300  
 Db 241 GAAATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGATCA 300  
 QY 301 GCGGTGAGCCCTTATCTCTACCTTCAATCCATCTCGATGTCGCTTGCATAGGCC 360  
 Db 301 GCGGTGAGCCCTTATCTCTACCTTCAATCCATCTCGATGTCGCTTGCATAGGCC 360  
 QY 361 ATAATGTTAAAGGTGTGATGAAGCAACCATCATTTGACATTTCTAACTAAGCGAAACAT 420

Db 361 ATAAATGGTTAAAGGTGTGGATGAAGCAACCATTCATTCACATTTCTAAGCGAAACAAT 420  
Qy 421 GCACAGCGTCAACAGATCAAGACGACATATCTCCAGGAACAGGAAGCCCTGGATGAA 480  
Db 421 GCACAGCGTCAACAGATCAAGACGACATATCTCCAGGAACAGGAAGCCCTGGATGAA 480  
Qy 481 ACATCTGAAGAAGCCCTTACAGGTCACTTGGAGAGTGTGTTTGTCTGTCTAAAAACT 540  
Db 481 ACATCTGAAGAAGCCCTTACAGGTCACTTGGAGAGTGTGTTTGTCTGTCTAAAAACT 540  
Qy 541 CGAGGCAATTTGATGCTGATGATGATCTGCTGCCATGAGGCGCTTGGAACTGATGAA 600  
Db 541 CGAGGCAATTTGATGCTGATGATGATCTGCTGCCATGAGGCGCTTGGAACTGATGAA 600  
Qy 601 GATACCTAAATGAGATTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGG 660  
Db 601 GATACCTAAATGAGATTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGG 660  
Qy 661 GTCTACAGAGAGAACTGAAGAGATCTGGCCAAAGACATTAACCTCAGACACATCTGGA 720  
Db 661 GTCTACAGAGAGAACTGAAGAGATCTGGCCAAAGACATTAACCTCAGACACATCTGGA 720  
Qy 721 GATTTTCGGAAGCGCTTGTCTTCTTCTTAAGGTGACCGATCTGAGGACTTTGGTGTG 780  
Db 721 GATTTTCGGAAGCGCTTGTCTTCTTCTTAAGGTGACCGATCTGAGGACTTTGGTGTG 780  
Qy 781 AATGAAGACTTGGCTGATTCAGATGCCAGGGGCTTGTATGAAGCAGGAGAAAGAGAA 840  
Db 781 AATGAAGACTTGGCTGATTCAGATGCCAGGGGCTTGTATGAAGCAGGAGAAAGAGAA 840  
Qy 841 GGGACAGAGCTAAACGCTGTTCAATACCATCTTACCACAGAGCTATCCAACTTCGC 900  
Db 841 GGGACAGAGCTAAACGCTGTTCAATACCATCTTACCACAGAGCTATCCAACTTCGC 900  
Qy 901 AGAGTGTTCAGAAATACACCAAGTACAGTAAAGTATGATGATGATGATGATGATGAT 960  
Db 901 AGAGTGTTCAGAAATACACCAAGTACAGTAAAGTATGATGATGATGATGATGATGAT 960  
Qy 961 GAGTTGAAAGGTGACATGAGAAATGCGCTCAGCTATCGTGAAGTGGCCACAGCBA 1020  
Db 961 GAGTTGAAAGGTGACATGAGAAATGCGCTCAGCTATCGTGAAGTGGCCACAGCBA 1020  
Qy 1021 CCAGCTTTCTTTTCAGAGAGCTTTCATCAAGCCATGAAAGGTGTTGGAATCGCCATA 1080  
Db 1021 CCAGCTTTCTTTTCAGAGAGCTTTCATCAAGCCATGAAAGGTGTTGGAATCGCCATA 1080  
Qy 1081 GCATTGATCAGGATTAATGTTTCCCGTTCTGAAATGATGATGATGATGATGATGATGAT 1140  
Db 1081 GCATTGATCAGGATTAATGTTTCCCGTTCTGAAATGATGATGATGATGATGATGATGAT 1140  
Qy 1141 TATCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
Db 1141 TATCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
Qy 1201 TATGAGAAATCTGCTGCTCTTTGAGGAACTAAACATTCCTTGTGTTGCTCA 1260  
Db 1201 TATGAGAAATCTGCTGCTCTTTGAGGAACTAAACATTCCTTGTGTTGCTCA 1260  
Qy 1261 GCTATGATCAGAGAGCTTTAAATATATATATATATATATATATATATATATATATAT 1320  
Db 1261 GCTATGATCAGAGAGCTTTAAATATATATATATATATATATATATATATATATATAT 1320  
Qy 1321 CTTCAACAGGATTAACAGTACCTACATGCTGAAATATAGCTTTTAAATCATTT 1380  
Db 1321 CTTCAACAGGATTAACAGTACCTACATGCTGAAATATAGCTTTTAAATCATTT 1380  
Qy 1381 TTATATTAACCTCTGTATATAGATTAAGTCCATTTTAAATATGTTTCCCAAC 1440  
Db 1381 TTATATTAACCTCTGTATATAGATTAAGTCCATTTTAAATATGTTTCCCAAC 1440  
Qy 1441 CATAAAACCCCTATACAGTGTGTTCTAGTAAACAATATACAGAAAGATGCTATGAGCTG 1500  
Db 1441 CATAAAACCCCTATACAGTGTGTTCTAGTAAACAATATACAGAAAGATGCTATGAGCTG 1500

Qy 1501 AAAATAAAATGNCGTC 1516  
Db 1501 AAAATAAAATGNCGTC 1516

RESULT 4  
US-09-864-864-321  
; Sequence 321, Application US/09864864  
; Patent No. US20020102679A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Dillon, Devin C.  
; APPLICANT: Secrist, Heather J.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Fling, Steve P.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Carter, Darick  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.523  
; CURRENT APPLICATION NUMBER: US/09/864,864  
; CURRENT FILING DATE: 2001-05-23  
; NUMBER OF SEQ ID NOS: 341  
; SOFTWARE: Corixa Invention Disclosure Database  
; SEQ ID NO 321  
; LENGTH: 1399  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-864-864-321

Query Match 91.6%; Score 1389.4; DB 9; Length 1399;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 125 AGTGTGAATCTTTCAGAGAGAAATTTCTCTTTAGTCTTTTGAAGAGGTAGAGATAAAG 184  
Db 1 AGTGTGAATCTTTCAGAGAGAAATTTCTCTTTAGTCTTTTGAAGAGGTAGAGATAAAG 60  
Qy 185 ACATTTTTCAGAAATGCGATGATCAGATTCCTCAAGCAGGCTGTTTATTGAAA 244  
Db 61 ACATTTTTCAGAAATGCGATGATCAGATTCCTCAAGCAGGCTGTTTATTGAAA 120  
Qy 245 ATGAAGAGCAGGAATATGTTTCAAACTGTGAAGTCATCCAAAGGTGGTCCGGATCAGCG 304  
Db 121 ATGAAGAGCAGGAATATGTTTCAAACTGTGAAGTCATCCAAAGGTGGTCCGGATCAGCG 180  
Qy 305 TGAGCCCTTATCTTACCTTCAATCCATCTCGATGCTGCTGCTTGCATAGGCCATAA 364  
Db 181 TGAGCCCTTATCTTACCTTCAATCCATCTCGATGCTGCTGCTTGCATAGGCCATAA 240  
Qy 365 TGGTTAAAGGTGGATGAAGCAACCATCATTCACATTTCTAACTAAGCGAAACAATGCAC 424  
Db 241 TGGTTAAAGGTGGATGAAGCAACCATCATTCACATTTCTAACTAAGCGAAACAATGCAC 300  
Qy 425 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGGAAGCCCTGGATGAAACAC 484  
Db 301 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGGAAGCCCTGGATGAAACAC 360  
Qy 485 TGAAGAAGCCCTTACAGGTCACTTGGAGAGGTTGTTTTAGTCTCTGCTAAAACTCCAG 544  
Db 361 TGAAGAAGCCCTTACAGGTCACTTGGAGAGGTTGTTTTAGTCTCTGCTAAAACTCCAG 420  
Qy 545 CGCAATTTGATGCTGATGAACCTTCTGCTGCCATGAAGGCCCTTGAACCTGATGAAGATA 604  
Db 421 CGCAATTTGATGCTGATGAACCTTCTGCTGCCATGAAGGCCCTTGAACCTGATGAAGATA 480  
Qy 605 CTCTAAATTTGATGATTTGGCATCAAGAACTCAAGAAATAACAGAAATCAGAGACATTAACAGGCTCT 664



725	QY	TTGCGAACCGTTTGGCTTTCTCTTGTCTAAGGGTGACCGATCTGAGGACTTTGGTGTGTAATG	784
601	DB	TTGCGAACCGTTTGGCTTTCTCTTGTCTAAGGGTGACCGATCTGAGGACTTTGGTGTGTAATG	660
785	QY	AGACATTGGCTGATTTCAGATGCCAGGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGGA	844
661	DB	AGACATTGGCTGATTTCAGATGCCAGGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGGA	720
845	QY	CAGACGCTAAACGCTGTTCAATATCCATCCTTACCAACGAGAGGCTATCCACAACCTTCGCAGAG	904
721	DB	CAGACGCTAAACGCTGTTCAATATCCATCCTTACCAACGAGAGGCTATCCACAACCTTCGCAGAG	780
905	QY	TGTTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTTGACCTCGAGT	964
781	DB	TGTTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTTGACCTCGAGT	840
965	QY	TGAAAGGTGCACATTGAGAAATGCCTCACAGCTATCGTGAAGTCGCGCCACAAGCAAAACCCAG	1024
841	DB	TGAAAGGTGCACATTGAGAAATGCCTCACAGCTATCGTGAAGTCGCGCCACAAGCAAAACCCAG	900
1025	QY	CTTTCTTTGCAGAGAGGCTTCATCAAGCCATGAAGAGGTGTGGAACCTCGCCATAAGGCAT	1084
901	DB	CTTTCTTTGCAGAGAGGCTTCATCAAGCCATGAAGAGGTGTGGAACCTCGCCATAAGGCAT	960
1085	QY	TGATCAGGATTATGGTTTCCCGTTCTGAAATTCACATGAATGATATCAAGCACTTCTATC	1144
961	DB	TGATCAGGATTATGGTTTCCCGTTCTGAAATTCACATGAATGATATCAAGCACTTCTATC	1020
1145	QY	AGAAGATGTATGGTATCTCCCTTTGCCAGCCATCCTGGATGAACCAAGGAGAGATG	1204
1021	DB	AGAAGATGTATGGTATCTCCCTTTGCCAGCCATCCTGGATGAACCAAGGAGAGATG	1080
1205	QY	AGAAATCCTGGTGGCTCTTTGTGGAGGAACCTAAACATTCCTCGATGGTCTCAAGCTA	1264
1081	DB	AGAAATCCTGGTGGCTCTTTGTGGAGGAACCTAAACATTCCTCGATGGTCTCAAGCTA	1140
1265	QY	TGATCAGAAGACTTTAATATATATTTTCATCCTATAAGCTTAAATAGGAAAGTTTCTTC	1324
1141	DB	TGATCAGAAGACTTTAATATATATTTTCATCCTATAAGCTTAAATAGGAAAGTTTCTTC	1200
1325	QY	AACAGGATTACAGTGTAGTACCTPACATGCTGAAAAATATAGCCTTTAAATCATTTTTAT	1384
1201	DB	AACAGGATTACAGTGTAGTACCTPACATGCTGAAAAATATAGCCTTTAAATCATTTTTAT	1260
1385	QY	ATTATRACTCTGATATATAGAGATAAGTCCATTTTTTAAAAATGTTTTTCCCCAAACCCATA	1444
1261	DB	ATTATRACTCTGATATATAGAGATAAGTCCATTTTTTAAAAATGTTTTTCCCCAAACCCATA	1320
1445	QY	AAAACCTTATACAAGTTGTTCTAGTAAACAATACATGAGAAAAAGATGCTATGTAGCTGAAAA	1504
1321	DB	AAAACCTTATACAAGTTGTTCTAGTAAACAATACATGAGAAAAAGATGCTATGTAGCTGAAAA	1380
1505	QY	TAAAAATGNCGTC	1516
1381	DB	TAAAAATGACGTC	1392

## RESULT 6

US-10-240-965-2

00 10 210 305 2  
; Sequence 2, Application US/10240965

Publication No. US20030165924A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: SHIFFMAN, DOV

APPLICANT: SOMOGYI, Roland

APPLICANT: LAWN, Richard M.

APPLICANT: SEILHAMER, Jeffrey J.

APPLICANT: PORTER, Gordon J

APPLICANT: MIKITA, Thomas

APPLICANT: TAI, Julie

TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION

FILE REFERENCE: PA-0025 PCT

FILE REFERENCE: FH-0023 FCI  
CURRENT APPLICATION NUMBER: US/10/240,965

QY 965 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCGTGAAGTGGCCACAAAGCAACCCAG 1024  
 DB 841 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCGTGAAGTGGCCACAAAGCAACCCAG 900  
 QY 1025 CTTTCTTTGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAACTGCGCATAGGCAT 1084  
 DB 901 CTTTCTTTGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAACTGCGCATAGGCAT 960  
 QY 1085 TGATCAGGATATGTTTCCCGTTCTGAAATGACATGAATGATATCAAAAGCAATTTCTATC 1144  
 DB 961 TGATCAGGATATGTTTCCCGTTCTGAAATGACATGAATGATATCAAAAGCAATTTCTATC 1020  
 QY 1145 AGAAGATGATGATATCTCCCTTTGCGAAGCCATCCTGGATGAACCAAGGAGAGATG 1204  
 DB 1021 AGAAGATGATGATATCTCCCTTTGCGAAGCCATCCTGGATGAACCAAGGAGAGATG 1080  
 QY 1205 AGAAATCTGCTGCTCTTTGAGAGAACTAAACATTCCTTGTGTTCTCAAGCTA 1264  
 DB 1081 AGAAATCTGCTGCTCTTTGAGAGAACTAAACATTCCTTGTGTTCTCAAGCTA 1140  
 QY 1265 TGATCAGAGACTTTAATATATATATTTTCATCCCTAATAGCTTAATAGGAAAGTTCTTC 1324  
 DB 1141 TGATCAGAGACTTTAATATATATATTTTCATCCCTAATAGCTTAATAGGAAAGTTCTTC 1200  
 QY 1325 RACAGGATACAGTGTAGCTACCTACATGCTGAAATATATAGCTTTAAATCATTTTAT 1384  
 DB 1201 AACAGGATACAGTGTAGCTACCTACATGCTGAAATATATAGCTTTAAATCATTTTAT 1260  
 QY 1385 ATTATAACTCTGTATAATAGAGATAAGTCCATTTTAAAAATGTTTCCCAAAACCAT 1444  
 DB 1261 ATTATAACTCTGTATAATAGAGATAAGTCCATTTTAAAAATGTTTCCCAAAACCAT 1320  
 QY 1445 AAACCTATACAGTGTGCTAGTAAACATACATGAGAAAGTCTATGAGCTGAAAA 1504  
 DB 1321 AAACCTATACAGTGTGCTAGTAAACATACATGAGAAAGTCTATGAGCTGAAAA 1380  
 QY 1505 TAAATGNCCTC 1516  
 DB 1381 TAAATGAGCTC 1392  
 RESULT 7  
 US-10-269-909-2  
 ; Sequence 2, Application US/10269909  
 ; Publication No. US20030180747A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HRUBAN, RALEH H.  
 ; APPLICANT: ARGANI, PEDRAM  
 ; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE  
 ; APPLICANT: MAITRA, ANIRBAN  
 ; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES  
 ; FILE REFERENCE: 58303(71699)  
 ; CURRENT APPLICATION NUMBER: US/10/269,909  
 ; CURRENT FILING DATE: 2003-10-11  
 ; PRIOR APPLICATION NUMBER: 60/328,609  
 ; PRIOR FILING DATE: 2001-10-11  
 ; PRIOR APPLICATION NUMBER: 60/332,754  
 ; PRIOR FILING DATE: 2001-11-19  
 ; NUMBER OF SEQ ID NOS: 87  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 1399  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-269-909-2  
 Query Match 91.5%; Score 1387.8; DB 14; Length 1399;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 125 AGTGTGAATCTTCAGAGAGAAATTTCTTTAGTGTCTTTGCAAGAGGTAGATAAAG 184

DB 1 AGTGTGAATCTTCAGAGAGAAATTTCTTTAGTGTCTTTGCAAGAGGTAGATAAAG 60  
 QY 185 ACATCTTTTCAAAAATGCGAATGGTATCAGAAATCCTCAAGCAGGCCCTGTTTATTTGAAA 244  
 DB 61 ACATCTTTTCAAAAATGCGAATGGTATCAGAAATCCTCAAGCAGGCCCTGTTTATTTGAAA 120  
 QY 245 ATGAAGAGCAGAGATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGATCAGCGG 304  
 DB 121 ATGAAGAGCAGAGATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGATCAGCGG 180  
 QY 305 TGAGCCCTTATCTTACCTTTCAATCCATCCTCGGATGTGCTGCTCCCTTGCATTAAGGCCATAA 364  
 DB 181 TGAGCCCTTATCTTACCTTTCAATCCATCCTCGGATGTGCTGCTCCCTTGCATTAAGGCCATAA 240  
 QY 365 TGGTTAAAGGTGTGATGAAGCAACCATCATTTCAATTTCTTAAGCGAAACAATGCAAC 424  
 DB 241 TGGTTAAAGGTGTGATGAAGCAACCATCATTTCAATTTCTTAAGCGAAACAATGCAAC 300  
 QY 425 AGCGTCAACAGATCAAAAGCAGCATATCTCCAGAAACAGGAAAGCCCTGGATGAACAC 484  
 DB 301 AGCGTCAACAGATCAAAAGCAGCATATCTCCAGAAACAGGAAAGCCCTGGATGAACAC 360  
 QY 485 TGAAGAAACCCCTTACAGGTCACTTGAAGAGGTGTTTGTAGCTCTGCTTAAATACTCCAG 544  
 DB 361 TTAAGAAAGCCCTTACAGGTCACTTGAAGAGGTGTTTGTAGCTCTGCTTAAATACTCCAG 420  
 QY 545 CGCAATTTGATGCTGATGAATCTTGTGCTGCCATGAAGGGCCTTGGAACTGATGAAGATA 604  
 DB 421 CGCAATTTGATGCTGATGAATCTTGTGCTGCCATGAAGGGCCTTGGAACTGATGAAGATA 480  
 QY 605 CTCATAATTGAGATTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGGCTCT 564  
 DB 481 CTCATAATTGAGATTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGGCTCT 540  
 QY 665 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAAACCTCAGACACATCTGGAGATT 724  
 DB 541 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAAACCTCAGACACATCTGGAGATT 600  
 QY 725 TTCGGAACGCTTTGCTTTCTTTGCTAAGGGTGAACCGATCTGAGGACTTTGGTGTGAATG 784  
 DB 601 TTCGGAACGCTTTGCTTTCTTTGCTAAGGGTGAACCGATCTGAGGACTTTGGTGTGAATG 660  
 QY 785 AAGACTTTGCTGATTCAGATGCCAGGGCCTTGTATGAAGCAGAGAGAAAGGAGGGA 844  
 DB 661 AAGACTTTGCTGATTCAGATGCCAGGGCCTTGTATGAAGCAGAGAGAAAGGAGGGA 720  
 QY 845 CAGACGTAAACGTTGTTCAATACCATCTTACCACCAAGAGCTATCCCACTTCGCGAG 904  
 DB 721 CAGACGTAAACGTTGTTCAATACCATCTTACCACCAAGAGCTATCCCACTTCGCGAG 780  
 QY 905 TGTTTCAGAAATACACCAAGTACAGTAAAGCATGACATGAACAAAGTTCTGGACCTGGAGT 964  
 DB 781 TGTTTCAGAAATACACCAAGTACAGTAAAGCATGACATGAACAAAGTTCTGGACCTGGAGT 840  
 QY 965 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCGTGAAGTGGCCCAACAAAGCAACCCAG 1024  
 DB 841 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCGTGAAGTGGCCCAACAAAGCAACCCAG 900  
 QY 1025 CTTTCTTTGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAACTGCGCATAGGCAT 1084  
 DB 901 CTTTCTTTGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAACTGCGCATAGGCAT 960  
 QY 1085 TGATCAGGATATGTTTCCCGTTCTGAAATGACATGAATGATATCAAAAGCAATTTCTATC 1144  
 DB 961 TGATCAGGATATGTTTCCCGTTCTGAAATGACATGAATGATATCAAAAGCAATTTCTATC 1020  
 QY 1145 AGAAGATGATGATATCTCCCTTTGCGAAGCCATCCTGGATGAACCAAGGAGAGATG 1204  
 DB 1021 AGAAGATGATGATATCTCCCTTTGCGAAGCCATCCTGGATGAACCAAGGAGAGATG 1080  
 QY 1205 AGAAATCTGCTGCTCTTTGAGAGAACTAAACATTTCCCTTGTGTTCTCAAGCTA 1264  
 DB 1081 AGAAATCTGCTGCTCTTTGAGAGAACTAAACATTTCCCTTGTGTTCTCAAGCTA 1140

QY 1265 TGATCAGAGACCTTTAATTATATATTTTCACTATAGCTTAAATAGGAAAGTTTCTTC 1324  
 DB 1141 TGATCAGAGACCTTTAATTATATATTTTCACTATAGCTTAAATAGGAAAGTTTCTTC 1200  
 QY 1325 AACAGGATACAGTGTAGCTACCTACATGCTGAAATATAGCTTTAAATCAATTTTAT 1384  
 DB 1201 AACAGGATACAGTGTAGCTACCTACATGCTGAAATATAGCTTTAAATCAATTTTAT 1260  
 QY 1385 ATTATAAATCTGTATATAGAGATAAGTCCATTTTAAATATGTTTCCCAAAACCATTA 1444  
 DB 1261 ATTATAAATCTGTATATAGAGATAAGTCCATTTTAAATATGTTTCCCAAAACCATTA 1320  
 QY 1445 AAACCCCTATCAAGTGTGTTCTTAGTAAATATACATGAGAAAGATGCTATGTAAGTGA 1504  
 DB 1321 AAACCCCTATCAAGTGTGTTCTTAGTAAATATACATGAGAAAGATGCTATGTAAGTGA 1380  
 QY 1505 TAAATGNCGTC 1516  
 DB 1381 TAAATGACGTC 1392

RESULT 8  
 US-10-133-937-2  
 ; Sequence 2, Application US/10133937  
 ; Publication No. US20030207278A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Khan, Javed  
 ; APPLICANT: Ringner, Markus  
 ; APPLICANT: Peterson, Carsten  
 ; APPLICANT: Meltzer, Paul  
 ; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,  
 ; TITLE OF INVENTION: DIAGNOSIS, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND  
 ; TITLE OF INVENTION: OTHER BIOLOGICAL STATES  
 ; FILE REFERENCE: 11613.56US01  
 ; CURRENT APPLICATION NUMBER: US/10/133,937  
 ; CURRENT FILING DATE: 2002-11-04  
 ; NUMBER OF SEQ ID NOS: 99  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 1399  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-133-937-2

Query Match 91.5%; Score 1387.8; DB 15; Length 1399;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 125 AGTGTGAATCTTCAGAGAGAAATTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAG 184  
 DB 1 AGTGTGAATCTTCAGAGAGAAATTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAG 60  
 QY 185 ACATTTTTCAAAATGCGCAATGGTATCAGAAATTCCTCAAGCAGGCGCTGGTTATTGAAA 244  
 DB 61 ACATTTTTCAAAATGCGCAATGGTATCAGAAATTCCTCAAGCAGGCGCTGGTTATTGAAA 120  
 QY 245 ATGAAGAGCGAGGAATATGTTTCAACTGTGGAAGTATCCTCAAGAGGTGTCCGGATCAGCGG 304  
 DB 121 ATGAAGAGCGAGGAATATGTTTCAACTGTGGAAGTATCCTCAAGAGGTGTCCGGATCAGCGG 180  
 QY 305 TGAGCCCTTATCTTACCTTCAATCCATCCCTCGGATGTCGCTTGCATTAAGGCCATAA 364  
 DB 181 TGAGCCCTTATCTTACCTTCAATCCATCCCTCGGATGTCGCTTGCATTAAGGCCATAA 240  
 QY 365 TGGTTAAAGGTGGATGAAGCAACCATATGTAATGTAATGTAATGTAATGTAATGTAATGTA 424  
 DB 241 TGGTTAAAGGTGGATGAAGCAACCATATGTAATGTAATGTAATGTAATGTAATGTAATGTA 300  
 QY 425 AGCTTCAACAGATCAAGAGAGCATATCTCCAGGAAACAGAGAGAGAGAGAGAGAGAGAGAG 484  
 DB 301 AGCTTCAACAGATCAAGAGAGCATATCTCCAGGAAACAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 485 TGAAGAAAGCCCTTACAGGTACCTTGAAGAGGTGTTTGTAGCTCTGCTAAAAAATCCAG 544  
 DB 361 TTAAGAAAGCCCTTACAGGTACCTTGAAGAGGTGTTTGTAGCTCTGCTAAAAAATCCAG 420  
 QY 545 CGCAATTTGATGCTGTGAATCTTCTGCTGCCATGAAGGSCCTTGGAACTGATGAAGATA 604  
 DB 421 CGCAATTTGATGCTGTGAATCTTCTGCTGCCATGAAGGSCCTTGGAACTGATGAAGATA 480  
 QY 605 CTCTAATTTGATGCTGTGAATCTTCTGCTGCCATGAAGGSCCTTGGAACTGATGAAGATA 664  
 DB 481 CTCTAATTTGATGCTGTGAATCTTCTGCTGCCATGAAGGSCCTTGGAACTGATGAAGATA 540  
 QY 665 ACAGAGAGGAACTGAGAGAGATCTGCCAAAGACATTAACCTCAGACACATCTCGAGATT 724  
 DB 541 ACAGAGAGGAACTGAGAGAGATCTGCCAAAGACATTAACCTCAGACACATCTCGAGATT 600  
 QY 725 TTCCGAACGCTTGTCTTCTTCTGCTAAGGCTGACCGATCTGAGGACTTTGGTGTGAATG 784  
 DB 601 TTCCGAACGCTTGTCTTCTTCTGCTAAGGCTGACCGATCTGAGGACTTTGGTGTGAATG 660  
 QY 785 AAGACTTGGCTGATTTCAGATGCCAGGSCCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA 844  
 DB 661 AAGACTTGGCTGATTTCAGATGCCAGGSCCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA 720  
 QY 845 CAGACGTAAAGCGTGTTCATATACCATCTTACCACCAAGAGCTATCCACAACTTTCGCAGAG 904  
 DB 721 CAGACGTAAAGCGTGTTCATATACCATCTTACCACCAAGAGCTATCCACAACTTTCGCAGAG 780  
 QY 905 TGTTTCAGAAATACACCAAGTACAGTAAGTACATGATGAATGATGATGATGATGATGATGATG 964  
 DB 781 TGTTTCAGAAATACACCAAGTACAGTAAGTACATGATGAATGATGATGATGATGATGATGATG 840  
 QY 965 TGAAGAGTGAATGAGAAATGCTCAGACGTATCGTGAAGTGCCTCAGCAAGCAAAACAG 1024  
 DB 841 TGAAGAGTGAATGAGAAATGCTCAGACGTATCGTGAAGTGCCTCAGCAAGCAAAACAG 900  
 QY 1025 CTTTCTTTGAGAGAGCTTCAACAGCCATGAAAGTGTGGAAGTGTGGAAGTGTGGAAGTGTG 1084  
 DB 901 CTTTCTTTGAGAGAGCTTCAACAGCCATGAAAGTGTGGAAGTGTGGAAGTGTGGAAGTGTG 960  
 QY 1085 TGATCAGGATTTATGTTTCCCGTTCTGAAATTTGACATGAAATGATATCAAAAGCATTTCTATC 1144  
 DB 961 TGATCAGGATTTATGTTTCCCGTTCTGAAATTTGACATGAAATGATATCAAAAGCATTTCTATC 1020  
 QY 1145 AGAAGATGATGTTATCTCCCTTTGCAAGCCATCTGGAATGAACCAAGAGGAGATATG 1204  
 DB 1021 AGAAGATGATGTTATCTCCCTTTGCAAGCCATCTGGAATGAACCAAGAGGAGATATG 1080  
 QY 1205 AGAAATCCCTGGTGGCTTTTGTGAGGAAATCAACATTTCCCTTGTATGCTCTCAAGCTA 1264  
 DB 1081 AGAAATCCCTGGTGGCTTTTGTGAGGAAATCAACATTTCCCTTGTATGCTCTCAAGCTA 1140  
 QY 1265 TGATCAGAGACTTTAATATATATTTTTCATTCCTATTAAGCTTAAATAGGAAAGTTTCTTC 1324  
 DB 1141 TGATCAGAGACTTTAATATATATTTTTCATTCCTATTAAGCTTAAATAGGAAAGTTTCTTC 1200  
 QY 1325 AACAGGATTAAGTGTAGCTACCTACATGCTGAAATATAGCTTTAAATCAATTTTAT 1384  
 DB 1201 AACAGGATTAAGTGTAGCTACCTACATGCTGAAATATAGCTTTAAATCAATTTTAT 1260  
 QY 1385 ATTATAAATCTGTATATAGAGATAAGTCCATTTTAAATATGTTTCCCAAAACCATTA 1444  
 DB 1261 ATTATAAATCTGTATATAGAGATAAGTCCATTTTAAATATGTTTCCCAAAACCATTA 1320  
 QY 1445 AAACCCCTATCAAGTGTGTTCTTAGTAAATATACATGAGAAAGATGCTATGTAAGTGA 1504  
 DB 1321 AAACCCCTATCAAGTGTGTTCTTAGTAAATATACATGAGAAAGATGCTATGTAAGTGA 1380  
 QY 1505 TAAATGNCGTC 1516  
 DB 1381 TAAATGACGTC 1392





Query Match 91.0%; Score 1379.4; DB 9; Length 1431;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1380; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 135 CTTACAGAGAAATTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAGACACATTTTC 194  
 DB 20 CTTACAGAGAAATTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAGACACATTTTC 79

QY 195 AAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCGCTGTTTATTTGAAAAATGAAGAGCA 254  
 DB 80 AAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCGCTGTTTATTTGAAAAATGAAGAGCA 139

QY 255 GGAATATGTTCAAACCTGTGAAGTATCCAAAGGTGTCCGGATCAGCGGTAGCGCCCTA 314  
 DB 140 GGAATATGTTCAAACCTGTGAAGTATCCAAAGGTGTCCGGATCAGCGGTAGCGCCCTA 199

QY 315 TCCTACCTTCAATCCATCCCTCGCATGTGCTGCTTGCATTAAGGCCCATTAATGTTAAAGG 374  
 DB 200 TCCTACCTTCAATCCATCCCTCGCATGTGCTGCTTGCATTAAGGCCCATTAATGTTAAAGG 259

QY 375 TGTGATGAAGCAACCATATTGACATTTCTAATGAAGCGAACAATGACAGCGTCACCA 434  
 DB 260 TGTGATGAAGCAACCATATTGACATTTCTAATGAAGCGAACAATGACAGCGTCACCA 319

QY 435 GATCAAGCAGCATATCTCCAGGAACAGGAAGCCCTGGATGAACACTGAAGAAAGC 494  
 DB 320 GATCAAGCAGCATATCTCCAGGAACAGGAAGCCCTGGATGAACACTGAAGAAAGC 379

QY 495 CTTACAGGTCACTTGAAGAGGTGTTTGTCTGTCTTAAATCCAGCGCAATTGA 554  
 DB 380 CTTACAGGTCACTTGAAGAGGTGTTTGTCTGTCTTAAATCCAGCGCAATTGA 439

QY 555 TGTGATGAAGCAACCATATTGACATTTCTAATGAAGCGAACAATGACAGCGTCACCA 614  
 DB 440 TGTGATGAAGCAACCATATTGACATTTCTAATGAAGCGAACAATGACAGCGTCACCA 499

QY 615 GATTTTGGCATCAAGAACTAACAAAGAAATTCAGAGACATTAACAGGGTCTACAGAGAGA 674  
 DB 500 GATTTTGGCATCAAGAACTAACAAAGAAATTCAGAGACATTAACAGGGTCTACAGAGAGA 559

QY 675 ACTGAAGAGAGATCTGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAAGCG 734  
 DB 560 ACTGAAGAGAGATCTGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAAGCG 619

QY 735 TTTGCTTTCTCTTCTAAGGGTGAACCATCTGAGGACTTTGTTGTTGAATGAAGACTTGGC 794  
 DB 620 TTTGCTTTCTCTTCTAAGGGTGAACCATCTGAGGACTTTGTTGTTGAATGAAGACTTGGC 679

QY 795 TGATTCAGATCCAGGGCTTGTATGAAGCAGGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 854  
 DB 680 TGATTCAGATCCAGGGCTTGTATGAAGCAGGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 739

QY 855 CGTGTTCATACCATCTTACCAAGAGGTATCCCAACTTCGACAGTGTTCAGAA 914  
 DB 740 CGTGTTCATACCATCTTACCAAGAGGTATCCCAACTTCGACAGTGTTCAGAA 799

QY 915 ATACACCAAGTACAGTAAAGCATGATGAACAAAGTTCTGACCTGAGTGTGAAGGTGA 974  
 DB 800 ATACACCAAGTACAGTAAAGCATGATGAACAAAGTTCTGACCTGAGTGTGAAGGTGA 859

QY 975 CATTCAGAAATGCTTCACAGCTATCTGTAAGTGGCCACAGCAAGCAACAGCTTTCTTTCG 1034  
 DB 860 CATTCAGAAATGCTTCACAGCTATCTGTAAGTGGCCACAGCAAGCAACAGCTTTCTTTCG 919

QY 1035 AGAAGATTCATCAAGCATGAAGGTGTGGAATCTCGGCATTAAGCATGATCAGGAT 1094  
 DB 920 AGAAGATTCATCAAGCATGAAGGTGTGGAATCTCGGCATTAAGCATGATCAGGAT 979

QY 1095 TATGTTTCCCGTCTCGAAATGATGAATGATCAAGCATTTCTATCAGAGAGATGA 1154  
 DB 980 TATGTTTCCCGTCTCGAAATGATGAATGATCAAGCATTTCTATCAGAGAGATGA 1039

RESULT 11  
 US-10-210-120-18  
 ; Sequence 18, Application US/10210120  
 ; Publication No. US20030175736A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chinnaiyan, Arul M.  
 ; APPLICANT: Rubin, Mark A.  
 ; APPLICANT: Sreekumar, Arun  
 ; TITLE OF INVENTION: Expression Profile of Prostate Cancer  
 ; FILE REFERENCE: UM-07221  
 ; CURRENT APPLICATION NUMBER: US/10/210,120  
 ; CURRENT FILING DATE: 2002-08-01  
 ; PRIOR APPLICATION NUMBER: US 60/309,581  
 ; PRIOR FILING DATE: 2001-08-02  
 ; PRIOR APPLICATION NUMBER: US 60/334,468  
 ; PRIOR FILING DATE: 2001-11-15  
 ; NUMBER OF SEQ ID NOS: 123  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 18  
 ; LENGTH: 1377  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-210-120-18

Query Match 90.2%; Score 1367.4; DB 14; Length 1377;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1368; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 147 ATTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAGACACATTTTCAAAAATGGCAAT 206  
 DB 1 ATTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAGACACATTTTCAAAAATGGCAAT 60

QY 207 GGTATCAGAAATTCCTCAAGCAGGCGCTGTTTATTTGAAAAATGAAGAGAGGAATATGTTCA 266  
 DB 61 GGTATCAGAAATTCCTCAAGCAGGCGCTGTTTATTTGAAAAATGAAGAGAGGAATATGTTCA 120

QY 267 AACTGTGAAGTCAATCAAGAGGTGTCCGGATCAGCGGTAGCGCCCTATCTACCTTCAA 326  
 DB 121 AACTGTGAAGTCAATCAAGAGGTGTCCGGATCAGCGGTAGCGCCCTATCTACCTTCAA 180

QY 327 TCCATCTCCGATGTCGCTGCTTGCATTAAGGCCATTAATGTTTAAAGGTGTGGATGAAGC 386  
 DB 181 TCCATCTCCGATGTCGCTGCTTGCATTAAGGCCATTAATGTTTAAAGGTGTGGATGAAGC 240

387 AACCATGATGACATTTCTAACTAAGCGAAACAATGACAGCGTCAACAGATCAAGCAGC 446  
Db  
241 AACCATGATGACATTTCTAACTAAGCGAAACAATGACAGCGTCAACAGATCAAGCAGC 300  
247 ATATCTCCAGGAAACAGGAAAGCCCTGGATGAAACAACTGAAGAAAGCCCTTACAGTCA 506  
Db  
301 ATATCTCCAGGAAACAGGAAAGCCCTGGATGAAACAACTGAAGAAAGCCCTTACAGTCA 360  
247 CTTTGAGAGGTTGTTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 566  
Db  
361 CTTTGAGAGGTTGTTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 420  
247 TCGTGCTGCCATGAGGAGCCCTTGAAGCTGTGATGATGATGATGATGATGATGATGATG 626  
Db  
421 TCGTGCTGCCATGAGGAGCCCTTGAAGCTGTGATGATGATGATGATGATGATGATGATG 480  
247 AAGAACTAACAAAGAAATCAGAGACATTAACAGGCTCTACAGAGAGAACTGAAGAGAGA 686  
Db  
481 AAGAACTAACAAAGAAATCAGAGACATTAACAGGCTCTACAGAGAGAACTGAAGAGAGA 540  
247 TCTGGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAAAGCTTTGCTTCTCT 746  
Db  
541 TCTGGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAAAGCTTTGCTTCTCT 600  
247 TGCTAAGGCTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTGATTCAGATG 806  
Db  
601 TGCTAAGGCTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTGATTCAGATG 660  
247 CAGGCTTGTATGAGCAG 866  
Db  
661 CAGGCTTGTATGAGCAG 720  
247 CATCTTTACCAACAG 926  
Db  
721 CATCTTTACCAACAG 780  
247 CAGTAAGCATGACATGAACAAAGTTCTGACCTGGAGTTGAAAGGTCAGATTCAGAAATG 986  
Db  
781 CAGTAAGCATGACATGAACAAAGTTCTGACCTGGAGTTGAAAGGTCAGATTCAGAAATG 840  
247 CCTCAGAGCTATGCTGAAGTGGCCCAAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1046  
Db  
841 CCTCAGAGCTATGCTGAAGTGGCCCAAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
247 TCAAGCCATGAAGGTTGTTGAACTCGGCATTAAGGCTTTCATCAGAGATTTATGTTTCCG 1106  
Db  
901 TCAAGCCATGAAGGTTGTTGAACTCGGCATTAAGGCTTTCATCAGAGATTTATGTTTCCG 960  
247 TTCTGAAATTCAGATGATGATATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1166  
Db  
961 TTCTGAAATTCAGATGATGATATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
247 TTGCAAGCCATCTCTGATGAAACCAAGAGAGAGATGATGAGAGAGAGAGAGAGAGAGAG 1226  
Db  
1021 TTGCAAGCCATCTCTGATGAAACCAAGAGAGAGATGATGAGAGAGAGAGAGAGAGAGAG 1080  
247 TGGAGAACTTAACATTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1286  
Db  
1081 TGGAGAACTTAACATTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
247 TATTTTTCATCTTAAAGCTTAAATAGGAAAGTTTCTTCAACAGAGATTCAGTGTAGCTAC 1346  
Db  
1141 TATTTTTCATCTTAAAGCTTAAATAGGAAAGTTTCTTCAACAGAGATTCAGTGTAGCTAC 1200  
247 CTACATGCTGAAATATAGCTTTAAATCAATTTTAAATCAATTTTAAATCAATTTTAAATCA 1406  
Db  
1201 CTACATGCTGAAATATAGCTTTAAATCAATTTTAAATCAATTTTAAATCAATTTTAAATCA 1260  
247 ATAAGTCCATTTTAAATATGTTTTCCTCAAAACCAATTAACCCCTTATCAAGATTTGTTCTA 1466  
Db  
1261 ATAAGTCCATTTTAAATATGTTTTCCTCAAAACCAATTAACCCCTTATCAAGATTTGTTCTA 1320  
247 GTAACATATCATGAGAAAGATGCTATGATGCTGAAATATAAATGTCGTC 1516

Db 1321 GTAACATATCATGAGAAAGATGCTATGATGCTGAAATATAAATGACGTC 1370  
RESULT 12  
US-09-917-800A-1585  
; Sequence 1585, Application US/09917800A  
; Patent No. US20020119462A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Castle, Arthur  
; APPLICANT: Blashoff, Michael  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5038-US  
; CURRENT APPLICATION NUMBER: US/09/917,800A  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,040  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,880  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: US 60/290,029  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US 60/290,645  
; PRIOR FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US 60/292,336  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/295,798  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/297,457  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,884  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,459  
; PRIOR FILING DATE: 2001-07-09  
; NUMBER OF SEQ ID NOS: 1740  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1585  
; LENGTH: 1402  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_012904  
US-09-917-800A-1585  
Query Match 58.5%; Score 886.8; DB 9; Length 1402;  
Best Local Similarity 80.1%; Pred. No. 3.9e-230;  
Matches 1101; Conservative 0; Mismatches 267; Indels 6; Gaps 5;  
QY 140 GAGAGAAATTTCTCTTTAGTTCTTTTGCAGA-AGGTAGAGATAAAGACACTTTTCAAAA 198  
Db 2 GAGCAAGCTTCTCTTCAGTTCCCTGGAGAGACAGCAATCAAGATACATTATTAATA 61  
QY 199 ATGGCAATGATCAGAAATTCCTCAAGCAGGCTGGTTTATTGAAAAATGAAGAGCAGAA 258  
Db 62 ATGGCAATGATCAGAAATTCCTCAAGCAGGCTGGTTTATTGAAAAATGAAGAGCAGAA 121  
QY 259 TATGTTCAAACTGTGAAGTATCCCAAGTGTCCGGATCAGCGGTGAGCCCTATCCT 318  
Db 122 TATGTTCAAGCTGTAAATTCCTACAAAGTGTCTCGATCAGCGGTGAGCCCTATCCT 181  
QY 319 ACCTTCAATCCATCTCGGATGTCTGCTGCTTGCATTAAGGCCATTAATGTTAAAGGTGTG 378  
Db 182 TCCTTCAATCCGCTCTCGGATGTCTGCTGCTTGCATTAAGGCCATTAATGTTAAAGGTGTG 241  
QY 379 GATGAAGCAACCATCATTTGACATTTCTTAACAGCAACCAATGACAGCGGTCAACAGATC 438  
Db 242 GATGAGCAACCATCATTTGACATTTCTTAACAGCAACCAATGACAGCGGTCAACAGATC 301  
QY 439 AAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAACACTGTAAGAGCCCTT 498

302 AAGGACGATACCTTACAGGAGACTGGAGAGCCCTGGATGAACCTTGAAAAAGCCCTT 361  
499 ACAGGTCACTTGAAGAGGTGTTTGTAGCTCTGCTAAATACTCCAGCGCAATTTGATGCT 558  
362 ACGGCGCACTGGAGAGGTGTTTGTGCTATGCTCAAGACCCAGCTCAGTTTATGCA 421  
559 GATGAATCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618  
422 GATGAATCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481  
619 TTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGAGGTCTACAGAGAGAACTG 678  
482 TTGACAAACAGATCTAACCCAGCAATCAGAGATTAAGAGTCTACAGAGAGAGCTG 541  
679 AAGAGAGTCTGCGCAAGACATAAAGCTCAGACACATCTGGAGATTTTCGGAACCTTTG 738  
542 AAAAGAGATCTGCGCAAGACATCACTTCGGACACATCTGGAGACTTTTCGTAATGCTT 601  
739 CTTTCTCTTGAAGGTGACCGATCTGAGGACTTTGCTGATGAAGACTTGGCTGAT 798  
602 CTTTCTCTGCGCAAGGTGATGCTGCTGAGGATTAAGTGTGAATCAAGATTTGCTGAT 661  
799 TCAGATGCCAGGCTTGTATGAAGCAGAGAAAGAGAAAGAGGAGACAGCTAAACGCTG 858  
662 ACAGATGCCAGGCTTGTATGAAGCTGAGAAAGAGAAAGAGGAGACAGCTGAAATGTG 721  
859 TTAATACCATCTTACCCAGAGCTATCCACACTTTCGAGAGTTCGAGAAATAC 918  
722 TTAATACCATCTTACCCAGAGCTATCCACACTTTCGAGAGTTCGAGAAATAC 781  
919 ACCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 978  
782 AGAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 841  
979 GAGAAATGCTCTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 1038  
842 GAGAAGTCTCTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 901  
1039 AAGCTTCAATCAAGCCATGAAGGTGTTGGAATCTCGCAATGAAGCAATGATCAGGATTA 1098  
902 AACTGTATGAAGCCATGAAGGTGTTGGAATCTCGCAATGAAGCAATGATCAGGATTA 961  
1099 GTTTCGGTCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1158  
962 GTCTCCGTTTCGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021  
1159 ATCTCCCTTTGCAAGGCTCTGGAATGAAACCAAGAGAGATGATGAGAAATCTCTGCTG 1218  
1022 ATCTCTCTGCAAGGCTCTGGAATGAAACCAAGAGAGATGATGAGAAATCTCTGCTG 1081  
1219 GCTCTTTGCGAGAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1277  
1082 GCTCTGCTGCGAGAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1141  
1278 TTAATATATATTTTCAATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1337  
1142 CTTAGCGTGTGTTTCTCTCTATGAGGCTTAAGTAGAAAGTTCCTTGTGCTAGTAA 1201  
1338 TGTAGTACCTACATGCTGAGAAATATAGCTTTTAAATCAATTTTATATATATATATAT 1397  
1202 TCTAATTAATCTTCTTCT 1259  
1398 ATATATAGATTA-AGTCCATTTTAAATGTTTTCCTCCCAACCAATTAACCTAT-AC 1455  
1360 ACAAATAGAGAAATATCTGTTTGTATATATATATATATATATATATATATATATAT 1319  
1456 AAGTTGTTCTAGTAAATACATGAGAAAGATGCTATGATGATGATGATGATGATGATGAT 1509  
1320 CAAGTCACTTTGTTACCATCTCTGAGAAAGAGTTTACATAGATAATAATAATAATA 1373

; Sequence 299, Application US/10316253  
; Publication No. US20030162706A1  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company  
; APPLICANT: Peters, Kevin  
; APPLICANT: Thompson, Larry  
; APPLICANT: Wang, Feng  
; APPLICANT: Greis, Kenneth  
; TITLE OF INVENTION: Angiogenesis Modulating Proteins  
; FILE REFERENCE: 8865M  
; CURRENT APPLICATION NUMBER: US/10/316,253  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/355,295  
; PRIOR FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 308  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 299  
; LENGTH: 1402  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (62)..(1102)  
; OTHER INFORMATION:  
; US-10-316-253-299

Query Match 58.5%; Score 886.8; DB 14; Length 1402;  
Best Local Similarity 80.1%; Pred. No. 3,9e-230;  
Matches 1101; Conservative 0; Mismatches 267; Indels 6; Gaps 5;

QY 140 GAGAAGAAATTCCTTTAGTCTTTGCAAGA-AGGTAGATGAAGACACTTTTTCAAAA 198  
DB 2 GAGCAAGCTTCTCTTCAGTTCCCTGGAAGACAGGCAATCAAGATATCTTTATTA 61  
QY 199 ATGGCAATGGTATCAGAAATTCCTCAAGCAGGCTGTTTATTGAAAAATGAAGACAG 258  
DB 62 ATGGCAATGGTATCAGAAATTCCTCAAGCAGGCTGTTTATTGAAAAATGAAGACAG 121  
QY 259 TATGTTCAAACTGTGAAGTCACTCAAAAGGTGTTCCCGATCAGGGTGAAGCCCTATCCT 318  
DB 122 TATGTTCAAGGTGTAATAATCCTCAAAAGGTGTTCCCGATCAGGGTGAAGCCCTATCCT 181  
QY 319 ACCTTCAATCCATCTCGGATGTCGCTGCTTGCATAAGGCCCATTAATGTTAAAGGTGT 378  
DB 182 TCCTTCAATCGTCTCGGATGTTGCTGCTTGCACAAAGCTATCATGTTTAAAGGTGT 241  
QY 379 GATGAAGCAACCATCATGTGATCTTAACTAAGGAAACAAATGCAAGCTCAAGATC 438  
DB 242 GATGAGCAACCATCATGTGATCTTAACTAAGGAAACCAATGCTCAGGCCAGCAGATC 301  
QY 439 AAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGATGAAACACCTGAAAGAAAGCCCTT 498  
DB 302 AAGCAGCATATCTTACAGGACATGCGAAGCCCTGATGAACCTTGAACCTTGAAGAAAGCCCTT 361  
QY 499 ACAGTCACTTGAAGAGGTGTTTATAGTCTGCTTAAAAAATCAGCGCAATTTGATGCT 558  
DB 362 ACGGCGCACTCGAGAGGTGTTTATAGTCTGCTTAAAAAATCAGCGCAATTTGATGCT 421  
QY 559 GATGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618  
DB 422 GATGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481  
QY 619 TTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGGCTCTACAGAGAGAACTG 678  
DB 482 TTGACAAACAGATCTAACCCAGCAATCAGAGATTAAGAGTCTACAGAGAGAGCTG 541  
QY 679 AAGAGAGTCTGCGCAAGACATAAAGCTCAGACACATCTGGAGATTTTCGGAACCTTTG 738  
DB 542 AAAAGAGATCTGCGCAAGACATCACTTCGGACACATCTGGAGACTTTTCGTAATGCTT 601  
QY 739 CTTTCTCTTGAAGGTGACCGATCTGAGGACTTTTCGAGACTTTTGTGAAATGAAGACTT 798  
DB 602 CTTTCTCTCGCAAGGTGATGCTGCTGAGGATATGATGATGATGATGATGATGATGATGAT 661



Query Match	39.4%;	Score	597.4;	DB	12;	Length	600;
Best Local Similarity	99.7%;	Pred. No.	1.1e-151;				
Matches	598;	Conservative	0;	Mismatches	2;	Indels	0;
Gaps	0;						
QY	906	GTTCAGAAATACACCAAGTACAGTACAGCATGACATGAACAAGTTCCTGACACTGGAGTT	965				
DB	1	GTTCAGAAATACACCAAGTACAGTACAGCATGACATGAACAAGTTCCTGACACTGGAGTT	60				
QY	966	GAAGAGTGCATTCAGAAATGCCCTCACAGCTATCGTGAAATGGGCCCAACAGCAAAACCCAGC	1025				
DB	61	GAAGGTGCATTCAGAAATGCCCTCACAGCTATCGTGAAATGGGCCCAACAGCAAAACCCAGC	120				
QY	1026	TTTCTTTGCAGAGAAGCTTCATCAAGCCATCAAGAGGTGTTGGAACTCGCCATAAGGCATT	1085				
DB	121	TTTCTTTGCAGAGAAGCTTCATCAAGCCATCAAGAGGTGTTGGAACTCGCCATAAGGCATT	180				
QY	1086	GATCAGGATTATGGTTTCCCGTCTCGAAATTCGATGAATCATCAAGCATCAAGCATTCATCA	1145				
DB	181	GATCAGGATTATGGTTTCCCGTCTCGAAATTCGATGAATCATCAAGCATTCATCA	240				
QY	1146	GAAGATGTATGGTATCTCCCTTTTGCCAAAGCCATCCTGGATCAAAACCAAGAGAGTATGA	1205				
DB	241	GAAGATGTATGGTATCTCCCTTTTGCCAAAGCCATCCTGGATCAAAACCAAGAGAGTATGA	300				
QY	1206	GAATACTCGTGGCTCTTTGTGGAGGAAACTTAACATTCCTTCATGGTCTCAAGCTAT	1265				



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

M nucleic - nucleic search, using sw model

run on: March 10, 2004, 09:18:38 / Search time 2509.9 Seconds

(without alignments)

18036.975 Million cell updates/sec

Title: US-10-084-817-8

Perfect score: 1516

Sequence: 1 cttgttttggacatagct.....gtcgaataaaatgncgtc 1516

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

EST:

1: em\_estba:\*

2: em\_esthm:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estc:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: em\_estfun:\*

15: em\_estom:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1030.6	68.0	1201	9	AL570884
C 2	1019	67.2	1169	9	AL570428
C 3	1014.4	66.9	1201	13	BX438944
4	1009.6	66.6	1128	13	BUS08007

5	1002	66.1	1183	9	AL544231
6	988.6	65.2	1163	13	BU902298
7	985.8	65.1	1201	9	AL553114
C 8	985.8	65.1	1201	9	AL575425
9	984	64.9	1201	9	AL541874
C 10	971.4	64.1	1136	9	AL574991
11	959.8	63.3	1201	9	AL553129
12	958.2	63.2	1201	9	AL544959
13	956	63.1	1164	9	AL551427
C 14	955.8	63.0	1201	9	AL576223
C 15	946.6	62.4	1201	9	AL553095
16	942	62.1	1076	12	BM476344
17	940.8	62.1	1077	12	BM473319
18	937.6	61.8	1201	9	AL551980
19	930.2	61.4	996	13	BQ222672
20	926.6	61.1	987	9	AL540651
C 21	925.2	61.0	1201	9	AL568911
22	910	60.0	1085	12	BM550642
23	906.4	59.8	1115	14	CK232619
24	903.6	59.6	990	13	BQ228162
C 25	901.4	59.5	1201	9	AL541873
26	897.4	59.2	1201	9	AL541443
27	894.2	59.0	1052	12	BM803761
28	893.8	59.0	1401	11	AK002681
29	891.4	58.8	1044	12	BM803483
30	880	58.0	939	12	BM468858
31	879.2	58.0	1059	14	CK232521
32	878.8	58.0	1201	9	AL542692
33	875.8	57.8	1186	9	AL553164
34	874.2	57.7	1033	12	BM460109
C 35	867.8	57.2	1181	9	AL576340
C 36	867.6	57.2	917	9	AL576603
C 37	866.4	57.2	973	9	AL568522
38	864.2	57.0	935	13	BQ959907
39	863	56.9	894	13	BU173401
40	857.8	56.6	880	13	BU902570
41	855.8	56.5	1054	12	BM458053
42	846.8	55.9	911	12	BM461778
43	840.4	55.4	942	9	AL553679
44	836	55.1	1044	12	BM555842
45	833	54.9	909	13	BU190139

## ALIGNMENTS

RESULT 1	AL570884/c	1201 bp	mRNA	linear	EST 31-MAY-2003
LOCUS	AL570884	Homo sapiens	PLACENTA COT 25-NORMALIZED	Homo sapiens	cdna
DEFINITION	clone CS0DI012YP05 3-PRIME, mRNA sequence.				
ACCESSION	AL570884				
VERSION	AL570884.2				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	On Feb 16, 2001 this sequence version replaced gi:12927628.				
Contact:	Genoscope				
	Genoscope - Centre National de Sequencage				
	BP 191 91006 Evry cedex - France				
	Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr				
	Library was constructed by Life Technologies, a division of				
	Invitrogen. This sequence belongs to sequence cluster 5436.f For				
	more information about this cluster, see				
	http://www.genoscope.cns.fr/				
	cgi-bin/cwuster.cgi?seq=CS0DI012CH03NP1&cluster=5436.f. Contact :				
	Feng Liang Email: fliang@lifetech.com URL :				



1060	AAATTAAGCAGCATWTCTCCA-GAAACACAGGAAGCCCCCTGGATGAAACACTCGAAGAAAG	1000
494	CCCTTACAGGTCACCTTGAGGAGGTTGTTTTAGCTCTGCTAAAAAATCCACGCGCAATTTG	553
1001	-CCYTACAGGTCACCTTGAGGAGGTTGTTTTAGCTCTGCTAAAAAATCCACGCGCAATTTG	943
554	ATGCTGATGAACCTTGTGCTGCCATGAAGGGCTTCGAACTGATGAAGATACCTCTAATTG	613
942	ATGCTGATGAACCTTGTGCTGCCATGAAGGGCTTGGMACTGATGAAGATACCTCTAATTG	893
614	AGATTTTGGCATCAAGAACTTAAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGG	673
882	AGATTTTGGCATCAAGAACTTAAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGG	823
674	AACCTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAGATTTTTCGGAACG	733
822	AACCTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAGATTTTTCGGAACG	763
734	CTTTTGCTTCTCTTTGCTTAAGGGTGAACGATCTCAGAGACTTTGGTGTGTAATGAAGACTTGG	793
762	CTTTTGCTTCTCTTTGCTTAAGGGTGACCGATCTGAGGACTTTGGTGTGTAATGAAGACTTGG	703
794	CTGATTCAGATCGCAGGGCCCTTGATGAAGCAGAGAAAGGAGAAAGGGGACAGACTTAA	853
702	CTGATTCAGATCGCAGGGCCCTTGATGAAGCAGAGAAAGGAGAAAGGGGACAGACTTAA	643
854	ACGTGTTCAATACCATCTTACACCAAGAGCTATCCACAACTTCGCAGAGTGTTCAGAA	913
642	ACGTGTTCAATACCATCTTACCAACCAAGAGCTATCCACAACTTCGCAGAGTGTTCAGAA	583
914	AATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTTGGAGTCTGAAAGGTG	973
582	AATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTTGGAGTCTGAAAGGTG	523
974	ACATTTGAGAAATGCTTCACAGCTATCTGTAAGTGGCGCCACAAAGAAACACAGCTTCTTTG	1033
522	ACATTTGAGAAATGCTTCACAGCTATCTGTAAGTGGCGCCACAAAGAAACACAGCTTCTTTG	463
1034	CAGAGAAGCTTCATCAAGCCATGAAGGTGTTGGAACTCGCCATTAAGGCATTCATCAGGA	1093
462	CAGAGAAGCTTCATCAAGCCATGAAGGTGTTGGAACTCGCCATTAAGGCATTCATCAGGA	403
1094	TTATGHTTCCCGTTCTGAAATTGACATGAATGATATCAAAAGCATTCCTATCAGAAGATGT	1153
402	TTATGHTTCCCGTTCTGAAATTAAACATGAATGAATCAAAAGCATTCCTATCAGAAGATGT	343
1154	ATGGTATCTCCCTTTGCGCAAGCCATCTCGATGAACCAAGAGAGAGATGAGAAAAATCC	1213
342	ATGGTATCTCCCTTTGCGCAAGCCATCTCGATGAACCAAGAGAGATGAGAAAAATCC	283
1214	TGTTGGCTCTTTTGTGGAGGAAAATAAAACATTCCTTGAATGGTCTCAGCTATGATCAGAA	1273
282	TGTTGGCTCTTTTGTGGAGGAAAATAAAACATTCCTTGAATGGTCTCAGCTATGATCAGAA	223
1274	GACTTTAAATTATATATTTTCACTCCTATAAGCTTAAATAGGAAAGTTTCTTCAACAGGATT	1333
222	GACTTTAAATTATATATTTTCACTCCTATAAGCTTAAATAGGAAAGTTTCTTCAACAGGATT	163
1334	ACAGTGTAGCTACCTACATGCTGAAAAAATATAGGCTTTAAATCATTTTTTATATTAACCT	1393
162	ACAGTGTAGCTACCTACATGCTGAAAAAATATAGGCTTTAAATCATTTTTTATATTAACCT	103
1394	CTGTATAATAGAGATAAGTCCATTTTTTAAAAATGTTTTCCCAAAACCAATAAAACCCCTAT	1453
102	CTGTATTATAGAGATAAGTCCATTTTTTAAAAATGTTTTCCCAAAACCAATAAAACCCCTAT	43
1454	ACAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTCTATGT	1495
42	ACAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTCTATGT	1

RESULT 3  
BX438944

LOCUS	BX438944	1201 bp	mRNA	linear	EST 15-MAY-2003
DEFINITION	BX438944 Homo sapiens PLACENTA Homo sapiens CDNA clone CS0DE004Y110				
	5-PRIME, mRNA sequence.				
ACCESSION	BX438944				
VERSION	BX438944.1	GI:30785738			
KEYWORDS	EST,				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Li, W.B., Gruber, C., Jesseee, J. and Polayes, D.				
JOURNAL	Full-length cDNA libraries and normalization				
COMMENT	Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 131 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 5436.f For more information about this cluster, see <a href="http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DE004BE05QP1&amp;cluster=5436.f">http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DE004BE05QP1&amp;cluster=5436.f</a> . Contact : Peng Liang Email : fliang@lifetech.com URL : <a href="http://fulllength.invitrogen.com/">http://fulllength.invitrogen.com/</a> Invitrogen Corporation 1600 Paraday Avenue Genoscope sequence ID : CS0DE004BE05QP1. Location/Qualifiers 1. 1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DE004Y110" /tissue_type="PLACENTA" /clone_lib="Homo sapiens PLACENTA" /notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, into double-strand cDNA was digested with Not I and cloned, into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."				
FEATURES	source				

## ORIGIN

	Query Match	66.9%;	Score 1014.4;	DB 13;	Length 1201;
	Best Local Similarity	94.7%;	Pred. No. 2.6e-259;		
	Matches 1036;	Conservative	22;	Mismatches 32;	Indels 4; Gaps 1;
Qy	142	GAAGAAATTTCTCTTTTAGTTCTCTTTGCAAGAGGTAGAGATAAGACACACTTTTTCAAAAATG	201		
Db	64	GAAGAAATTTCTCTTTTAGTTCTTTGCAAGAGGTAGAGATAAGACACACTTTTTCAAAAATG	123		
Qy	202	GCAATGGTATCAGAAATTCCTCAAGCAGGCGTGGTTTATTGAAAATGAAGAGCAGGGAATAT	261		
Db	124	GCAATGGTATCAGAAATTCCTCAAGCAGGCGTGGTTTATTGAAAATGAAGAGCAGGGAATAT	183		
Qy	262	GTTCCAACTGTGAAGTCATCCAAAGGTGGTCCCGGATCAGCGGTGAGCCCTATCCTPACC	321		
Db	184	GTTCCAACTGTGAAGTCATCCAAAGGTGGTCCCGGATCAGCGGTGAGCCCTATCCTPACC	243		
Qy	322	TTCAATCCATCCTCGGATGTCGCTGCCCTTCGATAAGGCCATAATGGTTAAAGGTGTGGAT	381		
Db	244	TTCAATCCATCCTCGGATGTCGCTGCCCTTCGATAAGGCCATAATGGTTAAAGGTGTGGAT	303		
Qy	382	GAAGCAACCATCATTTGACATTTCTAACTTAAGCGHAAACAATGCACAGCGTCAACAGATCAAA	441		
Db	304	GAAGCAACCATCATTTGACATTTCTAACTTAAGCGHAAACAATGCACAGCGTCAACAGATCAAA	363		
Qy	442	GCAGCATATCTCCAGGAAAACAGGAAAGCCCTCGATGAAAACACTGGAAGAAAGCCCTTACA	501		
Db	364	GCAGCATATCTCCAGGAAAACAGGAAAGCCCTCGATGAAAACACTGGAAGAAAGCCCTTACA	423		
Qy	502	GGTCACTTTCAGGAGGTGTTTTAGTCTCTGCTAAAAACTCAGCGCAATTTTGATGCTGTAT	561		
Db	424	GGTCACTTTCAGGAGGTGTTTTAGTCTCTGCTAAAAACTCAGCGCAATTTTGATGCTGTAT	483		

```
QY 562 GAACTTCGTCTGCCATGAAGAGCCCTTGGAACTGATGAAGATACTCTAATTGAGATTTTG 621
D 484 GAACTTCGTCTGCCATGAAGAGCCCTTGGAACTGATGAAGATACTCTAATTGAGATTTTG 543
QY 622 GCATCAGAACTCAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGAACTGAAG 581
D 544 GCATCAGAACTCAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGAACTGAAG 503
QY 682 AGAGATCTGCCAAAGACATAAAGCTCTGAGACATCTGGAGATTTTCGGAAACGCTTTGCTT 741
D 604 AGAGATCTGCCAAAGACATAAAGCTCTGAGACATCTGGAGATTTTCGGAAACGCTTTGCTT 663
QY 742 TCTCTCTTAAAGGTGACCATCTGAGGACTTTGGTGTGATGAAGACATCTGGTGTATCA 801
D 664 TCTCTCTTAAAGGTGACCATCTGAGGACTTTGGTGTGATGAAGACATCTGGTGTATCA 723
QY 802 GATGCCAGGGCCCTTGTATGAAGCAGAGAGAAAGGAGAAAGGGGACAGACGTAAACGTGTTTC 861
D 724 GATGCCAGGGCCCTTGTATGAAGCAGAGAGAAAGGAGAAAGGGGACAGACGTAAACGTGTTTC 783
QY 862 AATACCATCTTACCACAGAGAGCTATCCAACTCTCGACCTTGGAGTTGAAAGGTGACATTTGAG 921
D 784 AATACCATCTTACCACAGAGAGCTATCCAACTCTCGACCTTGGAGTTGAAAGGTGACATTTGAG 843
QY 922 AAGTACAGTAAAGATGACATGAACAAAGTTCTGGACCTTGGAGTTGAAAGGTGACATTTGAG 981
D 844 AAGTACAGTAAAGATGACATGAACAAAGTTCTGGACCTTGGAGTTGAAAGGTGACATTTGAG 903
QY 982 AATGCTCTCAGATATCGGTGAAGTGGCCCAAGCAACACAGCTTTCTTTGGCAGAGAA 1041
D 904 AATGCTCTCAGATATCGGTGAAGTGGCCCAAGCAACACAGCTTTCTTTGGCAGAGAA 963
QY 1042 CTTTCATCAAGCCATGAAGGTGTTGGAATCGGCAATGAGCATTTGATCAGGATTTAGTT 1101
D 964 CTTTCATCAAGCCATGAAGGTGTTGGAATCGGCAATGAGCATTTGATCAGGATTTAGTT 1023
QY 1102 TCCGTTTCTGAAATGATGAATGATATCAAGCATTTCTATCAGAAAGATGATGATGATC 1161
D 1024 TCCGTTTCTGAAATGATGAATGATATCAAGCATTTCTATCAGAAAGATGATGATGATGAT 1083
QY 1162 TCCGTTTCTGAAATGATGAATGATATCAAGCATTTCTATCAGAAAGATGATGATGATGAT 1221
D 1084 CCGTTTCCACCACTCTGAGTGAAGCAAGCAAGGAGATGATGAGAAATCTCTGTTGGCT 1139
QY 1222 CTTTGTGGAGGAAA 1235
D 1140 TTTGGGGGRTAAA 1153
```

```
RESULT 4
BU508007
LOCUS
DEFINITION BU508007 1128 bp mRNA linear EST 12-SEP-2002
5', mRNA sequence.
AGENCOURT 10128424 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6502033
ACCESSION BU508007
VERSION BU508007.1 GI:22814240
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1128)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
```

```
http://image.llnl.gov
Plate: L14M14057 Row: h Column: 02
High quality sequence stop: 828.
FEATURES
Location/Qualifiers
source 1..1128
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6502033"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 Kb. "
```

## ORIGIN

```
Query Match 66.6%; Score 1009.6; DB 13; Length 1128;
Best Local Similarity 97.8%; Pred. No. 4.9e-258;
Matches 1076; Conservative 0; Mismatches 19; Indels 5; Gaps 5;
QY 145 GAATTTCTCTTAGTCTTTTTCGAAAGAGGTAGAGATAAGACA-CTTTTCAAAAATGGC 203
D 1 GAATTTCTCTTAGTCTTTTTCGAAAGAGGTAGAGATAAGACANCTTTTCAAAAATGGC 60
QY 204 AATGGTATCAGAAATTCCTCAAGCAGGCTGCTTTATTGAAAATGAAGAGCAGGAATATGT 263
D 61 AATGGTATCAGAAATTCCTCAAGCAGGCTGCTTTATTGAAAATGAAGAGCAGGAATATGT 120
QY 264 TCAAACTGTGAAGTCATCCAAAAGGTGTCGGATCAGCGGTGAGCCCTTATCTACTCTT 323
D 121 TCAAACTGTGAAGTCATCCAAAAGGTGTCGGATCAGCGGTGAGCCCTTATCTACTCTT 180
QY 324 CAATCCATCTCGGATCTCGCTTGCATTAAGCCCATTAATGGTTAAAGGTGTTGGATGA 383
D 181 CAATCCATCTCGGATCTCGCTTGCATTAAGCCCATTAATGGTTAAAGGTGTTGGATGA 240
QY 384 AGCAACCATCATTCACATTTCTAATAAGCGAAACAAATGCAAGCGTCAACAGATCAAAAGC 443
D 241 AGCAACCATCATTCACATTTCTAATAAGCGAAACAAATGCAAGCGTCAACAGATCAAAAGC 300
QY 444 AGCATATCTCCAGAGAAACAGAGAGCCCTTGGATGAACACACTGAAGAAAGCCCTTACAGG 503
D 301 AGCATATCTCCAGAGAAACAGAGAGCCCTTGGATGAACACACTGAAGAAAGCCCTTACAGG 360
QY 504 TCACCTTGAGGAGGTGTTTGTAGCTCTGTAAATACTCCAGCGCAATTTGATGCTGATGA 563
D 361 TCACCTTGAGGAGGTGTTTGTAGCTCTGTAAATACTCCAGCGCAATTTGATGCTGATGA 420
QY 564 ACTTCGTCTGCCATGAAGGCTTGGAACTGATGAAGATACTCTAATTGAGATTTTGGC 623
D 421 ACTTCGTCTGCCATGAAGGCTTGGAACTGATGAAGATACTCTAATTGAGATTTTGGC 480
QY 624 ATCAAGAACTAACAAAGAAATCAGAGACATTTAAACAGGGTCTACAGAGAGAACTGAAGAG 683
D 481 ATCAAGAACTAACAAAGAAATCAGAGACATTTAAACAGGGTCTACAGAGAGAACTGAAGAG 540
QY 684 AGATCTGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAAACGCTTTGCTTTC 743
D 541 AGATCTGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAAACGCTTTGCTTTC 600
QY 744 TCTTGTCTAAGGGTGACCGATCTGAGGACTTTTGGTGTGAATGAAGACTTGGCTGATTCAGA 803
D 601 TCTTGTCTAAGGGTGACCGATCTGAGGACTTTTGGTGTGAATGAAGACTTGGCTGATTCAGA 660
QY 804 TCCAGGGCCCTTGTATGAGCAGAGAGAAAGGAGAAAGGGGACAGACCTTAACGTTTCAA 863
D 661 TCCAGGGCCCTTGTATGAGCAGAGAGAAAGGAGAAAGGGGACAGACCTTAACGTTTCAA 720
QY 864 TACCATCTTACCACCAAGAGCTATCCACAACTTCGACAGAGTGTTCAGAAATACACAA 923
D 721 TACCATCTTACCACCAAGAGCTATCCACAACTTCGACAGAGTGTTCAGAAATACACAA 780
```

QY 924 GTACAGTAGCATGACATGACAAAGTTCTGACCTGGAGTTGAAGTGACATTGAGAA 983  
 DB 781 GTACAGTAGCATGACATGACAAAGTTCTGACCTGGAGTTGAAGTGACATTGAGAA 840  
 QY 984 ATGCGCTCACAGCTATCGTGAAGTGGCCACACAGCAAAACAGCTTTCTTTGACAGAGAGCT 1043  
 DB 841 ATGCGCTCACAGCTATCGTGAAGTGGCCACACAGCAAAACAGCTTTCTTTGACAGAGAGCT 900  
 QY 1044 TCATCAGCCATGAAGGTTGGAACTCGCCATGAAGCATGATGACAGATAT-GGTTT 1102  
 DB 901 TCATC-AGCCATGAAGGTTGGAACTCGCCATGAAGCATGATGACAGATATGGGTTT 959  
 QY 1103 CCGGTTCTGAATGACATGAATGATATCAAAAGCATTTCTATCAGAGATGATGATGATCT 1162  
 DB 960 CCGGTTCTGAATGACATGAATGATATCAAAAGCATTTCTATCAGAGATGATGATGATCT 1019  
 QY 1163 -CCGTTGCGAAGCCATCTCGATGAAGCAAAAGAGAGATGATGAGAAAATCTGCTGG-GGC 1220  
 DB 1020 CCGCTTTGCGAAGCCATCTCGGATGAAGCAAAAGAGAGATTTTGTAAAATCTGCTGGGC 1079  
 QY 1221 TCTTTGCGAAGCAAACTAA 1240  
 DB 1080 TCTTTGCGAAGCAAACTAA 1099

RESULT 5  
 AL544231 1183 bp mRNA linear EST 31-MAY-2003  
 LOCUS AL544231 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS0D1019F12 5-PRIME, mRNA sequence.

ACCESSION AL544231  
 VERSION AL544231.2 GI:31266076  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1183)  
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:12876711.  
 CONTACT Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 5436.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0D1019DC06QPL&cluster=5436.f. Contact :  
 Feng Liang Email: fliang@lifetech.com URL: Corporation 1600  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID: CS0D1019DC06QPL.  
 Location/Qualifiers

FEATURES  
 source  
 1. .1183  
 /organism="Homo sapiens"  
 /mol\_type="RNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1019F12"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 66.1%; Score 1002; DB 9; Length 1183;  
 Best Local Similarity 95.3%; Pred. No. 5.3e-256;  
 Matches 1050; Conservative 13; Mismatches 33; Indels 6; Gaps 3;  
 QY 136 TTACAGAGAGAAATTCCTTAGTCTTTGCAAGAGGTAGAGATAAGACACTTTTCA 195

DB 45 TCCGGGATGAATTTCTTTAGTTCTTTGCAAGAGGTAGAGATAAGACACTTTTCA 104  
 QY 196 AAAATGCAATGATATCAGAAATTCCTCAAGCAGGCTGTTTATTGAAAATGAAGAGCAG 255  
 DB 105 AAAATGCAATGATATCAGAAATTCCTCAAGCAGGCTGTTTATTGAAAATGAAGAGCAG 164  
 QY 256 GAATATGTTCAAACTCTGAAGTCAATCAAAAGTGGTCCGGATCAGCGGTGAGCCCTAT 315  
 DB 165 GAATATGTTCAAACTCTGAAGTCAATCAAAAGTGGTCCGGATCAGCGGTGAGCCCTAT 224  
 QY 316 CCTACCTTCAATCCATCTCGGATGTCGTCCTTCATCAAGGCCATTAATGGTTAAAGGT 375  
 DB 225 CCTACCTTCAATCCATCTCGGATGTCGTCCTTCATCAAGGCCATTAATGGTTAAAGGT 284  
 QY 376 GTGATGACAGCAACCATCTTGACATTTAACTTAAGCGAAACAATCAGACGCTCAACAG 435  
 DB 285 GTGATGACAGCAACCATCTTGACATTTAACTTAAGCGAAACAATCAGACGCTCAACAG 344  
 QY 436 ATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTCGATGAAACACCTGAAGAAGCC 495  
 DB 345 ATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTCGATGAAACACCTGAAGAAGCC 404  
 QY 496 CTTACAGGTCACCTTGAGGAGGTTGTTTGTAGCTCTCTAAAPACTCCAGCGCAATTTGAT 555  
 DB 405 CTTACAGGTCACCTTGAGGAGGTTGTTTGTAGCTCTCTAAAPACTCCAGCGCAATTTGAT 464  
 QY 556 GCTGATGAATCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615  
 DB 465 GCTGATGAATCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 524  
 QY 616 ATTTTGGCATCAAGMACTAACAAGAAATCAGAGACATTAACACAGGCTCTACAGAGGAA 675  
 DB 525 ATTTTGGCATCAAGMACTAACAAGAAATCAGAGACATTAACACAGGCTCTACAGAGGAA 584  
 QY 676 CTGAAGAGAGATCTGCGCAAGACATAAACCCTCAGACACATCTCGAGATTTTCGGAAGCT 735  
 DB 585 CTGAAGAGAGATCTGCGCAAGACATAAACCCTCAGACACATCTCGAGATTTTCGGAAGCT 644  
 QY 736 TTGCTTTCTCTGCTAAGGCTGACCGATCTGAGGACTTTGGTGTGATGAGAGACTTGGCT 795  
 DB 645 TTGCTTTCTCTGCTAAGGCTGACCGATCTGAGGACTTTGGTGTGATGAGAGACTTGGCT 704  
 QY 796 GATTACAGATCGCAGGCTTTGTATGAAGCAGGAGAAAGGAGAAAGGAGCAGACGCTAAAC 855  
 DB 705 GATTACAGATCGCAGGCTTTGTATGAAGCAGGAGAAAGGAGAAAGGAGCAGACGCTAAAC 764  
 QY 856 GTGTTCAATACCATCTCTTACCAAGAGTATCCCACTTCGAGAGTGTTCAGAAA 915  
 DB 765 GTGTTCAATACCATCTCTTACCAAGAGTATCCCACTTCGAGAGTGTTCAGAAA 824  
 QY 916 TACACCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 975  
 DB 825 TACACCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 884  
 QY 976 ATTGAGAAATGCGCTCAGCTATCGTGAAGTGGCCCAAGAGAAACAGGCTTTCTTTGCA 1035  
 DB 885 ATTGAGAAATGCGCTCAGCTATCGTGAAGTGGCCCAAGAGAAACAGGCTTTCTTTGCA 943  
 QY 1036 GAGAGCTTCTCAAGCCATGAAGGTTGTGAACTCGCCATAGGCAATGATGATGATGATGAT 1095  
 DB 944 GAGAGCTTCTCAAGCCATGAAGGTTGTGAACTCGCCATAGGCAATGATGATGATGATGAT 1003  
 QY 1096 ATGTTTCCCGTCTCGAAATTTGACATGAATGATATCAAAAGCATTTCTATCAGAGATGAT 1155  
 DB 1004 ATGTTTCCCGTCTCGAAATTTGACATGAATGATATCAAAAGCATTTCTATCAGAGG--GTA 1061  
 QY 1156 GGTATCTCCCTTGGCAAGCCATCTCGATGAACCAAGAGAGATGATGAGAAATCTCTG 1215  
 DB 1062 TGGWYTCCTTTGGCAAGCCWYCTGG---ATGAACCAAGAGAGATGATGAGAAATCTCTG 1118  
 QY 1216 GTGGCTCTTTGTTGGAGGAAACT 1237

Db 1119 GGCTCTTGTGGGAAATTAAMT 1140

RESULT 6  
BU902298  
LOCUS  
DEFINITION AGENCOURT\_10127317 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:550136  
5' mRNA sequence.  
ACCESSION BU902298  
VERSION BU902298.1 GI:24084211  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1163)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
Cloning Strategy: Agencourt Bioscience Corporation  
Cloning Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L14M14050 Row: f Column: 01  
High quality sequence stop: 821.  
Location/Qualifiers  
1. 1163  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6503136"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dh. Average insert size 2.1 Kb."  
ORIGIN  
Query Match 65.2%; Score 988.6; DB 13; Length 1163;  
Best Local Similarity 95.2%; Pred. No. 2e-252;  
Matches 1063; Conservative 0; Mismatches 49; Indels 5; Gaps 4;  
135 CTTGAGAGAGAAATTTCTTTAGTCTTTGCAAGAGAGGTAGAGATAAAGACA-CTTTT 193  
24 CTTGAGAGAGAAATTTCTTTAGTCTTTGCAAGAGAGGTAGAGATAAAGACATTTT 83  
194 CAAAAATGGCAATGGTATCAGAAATCTCTCAAGAGGCGCTGTTTATTGAAATGAGAGC 253  
84 CAAAAATGGCAATGGTATCAGAAATCTCTCAAGAGGCGCTGTTTATTGAAATGAGAGC 143  
254 AGGAATATGTTCAAACTGTAAGTATCATCAAGAGTGTCCGGATCAGCGGTGAGCCCT 313  
144 AGGAATATGTTCAAACTGTAAGTATCATCAAGAGTGTCCGGATCAGCGGTGAGCCCT 203  
314 ATCTTACCTTCAATCCATCCGATGCGCTGCTTGCATTAAGGCCATTAATGGTTAAAG 373  
204 ATCTTACCTTCAATCCATCCGATGCGCTGCTTGCATTAAGGCCATTAATGGTTAAAG 263  
374 GTGTGGATGAGCAACCATTCACATTCCTAACTAAGCGAAACAAATGCACAGCTCAAC 433  
264 GTGTGGATGAGCAACCATTCACATTCCTAACTAAGCGAAACAAATGCACAGCTCAAC 323  
434 AGATCAAAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAAACACTGAGAAAG 493  
324 AGATCAAAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAAACACTGAGAAAG 383  
494 CCCCTACAGGTACCTTGGAGAGGTTGTTTAGCTCTGCTTAAAACTCCAGCGCAATTTG 553

Db 384 CCCTTACAGGTACCTTGGAGAGTGTGTTTGTCTGCTGCTTAAAACTCCAGCGCAATTTG 443

Qy 554 ATGCTGATGAATCTTGTCTGCCATGAAGGCGCTTGGAACTGATGAAGATACTCTAAATTTG 613

Db 444 ATGCTGATGAATCTTGTCTGCCATGAAGGCGCTTGGAACTGATGAAGATACTCTAAATTTG 503

Qy 614 AGATTTTGGCATCAAGACTAAACAAAGAAATCAGAGACATTAACAGGGTCTTACAGAGAGG 673

Db 504 AGATTTTGGCATCAAGACTAAACAAAGAAATCAGAGACATTAACAGGGTCTTACAGAGAGG 563

Qy 674 AACTGAGAGAGATCTGGCCAAAGACATACCTCAGACACATCTGGAGATTTTCGGAAGC 733

Db 564 AACTGAGAGAGATCTGGCCAAAGACATACCTCAGACACATCTGGAGATTTTCGGAAGC 623

Qy 734 CTTTGTCTTCTTCTTGAAGGTTGACGATCTGAGGCTTGGTGTGAATGAAGACTTTGG 793

Db 624 CTTTGTCTTCTTCTTGAAGGTTGACGATCTGAGGCTTGGTGTGAATGAAGACTTTGG 683

Qy 794 CTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGGACAGCGTAA 853

Db 684 CTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGGACAGCGTAA 743

Qy 854 AGTGTTCATACCATCTTACCCAGAGAGCTATCCACACTTCCGAGAGTGTTCAGA 913

Db 744 AGTGTTCATACCATCTTACCCAGAGAGCTATCCACACTTCCGAGAGTGTTCAGA 803

Qy 914 AATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGTG 973

Db 804 AATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGTG 863

Qy 974 ACATTGAGAAATGCTCAGCTATCAGCTATCGTGAAGTGGCCACCAAGCAACACCGCTTTCTTTG 1033

Db 864 ACATTGAGAAATGCTCAGCTATCAGCTATCGTGAAGTGGCCACCAAGCAACACCGCTTTCTCTTG 923

Qy 1034 CAGAGAAGCTTCATCAAGCCATGAAGGTTGTGAAGTTCGCACTCGCCATAAGGCATTTGATCAGGA 1093

Db 924 CAGAGAAGCTTCATCAAGCCATGAAGGTTGTGAAGTTCGCACTCGCCATAAGGCATTTGATCAGGG 983

Qy 1094 TTATGGTTTCCGGTCTGAAATGACATGAATGATAT-CAAAGCATTTCTATCAGAG-AT 1151

Db 984 ATATGGTTTCCGGTCTGAAATGACATTTGAATGAATCAAGGCTTTCTATCAAAAGAG 1043

Qy 1152 GPATGATATCTCCCTTTGCCAAGCCATCCTCGAT--GAAACCAAGGAGAGATGATGAGAAA 1209

Db 1044 GAATGAAATCTCCCTTGGCACACCCCTCTGTTGTAACCAAGGGAATTAATGGAATA 1103

Qy 1210 ATCTCTGGTCTTCTTGTGGAGGAACTAATCAATTC 1246

Db 1104 AATCCCGGGGCGCTCTTTGTGGGAAGAAAAAATAC 1140

RESULT 7  
AL553114 1201 bp mRNA linear EST 31-MAY-2003  
LOCUS AL553114 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CS01074YC03 5-PRIME, mRNA sequence.  
ACCESSION AL553114  
VERSION AL553114.2 GI:31274928  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1201)  
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On Feb 15, 2001 this sequence version replaced gi:12892649.  
Contact: Genoscope  
Genoscope - Centre National de Sequenage  
Bp 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr



Db 1022 AAAAGCCCTMACAGGTCMCCTTGAGGAGGTTGTTTAAKGTCTGCTAAACACTCCAGCGC 963

QY 548 AATTGTGATGATGAACCTTCGTGCTGCCATGAAGGCGCTTGGAACCTGATGAAGATGCTC 607  
 Db 962 AATTGTGATGATGAACCTTCGTGCTGCCATGAAGGCGCTTGGAACCTGATGAAGATGCTC 903  
 QY 608 TAATTGAGATTTGGCATCAAGATTAACAAAGAAATCAAGACATTAACAGAGGCTTCA 667  
 Db 902 TAATTGAGATTTGGCATCAAGATTAACAAAGAAATCAAGACATTAACAGAGGCTTCA 843  
 QY 668 GAGAGGAACCTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAGATTTTC 727  
 Db 842 GAGAGGACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAGATTTTC 783  
 QY 728 GGAACGCTTTGCTTCTTCCTTAAGGCTGACGATGAGGACTTTGGTGTGAATGAAG 787  
 Db 782 GGAACGCTTTGCTTCTTCCTTAAGGCTGACGATGAGGACTTTGGTGTGAATGAAG 723  
 QY 788 ACTTGGCTGATTCAGATGCGAGGCGCTTGTATGAAGCAGAGAGAAAGAGAGGAGACAG 847  
 Db 722 ACTTGGCTGATTCAGATGCGAGGCGCTTGTATGAAGCAGAGAGAAAGAGAGGAGACAG 863  
 QY 848 ACGTAAACGCTGCAATACCATCTTACCACCAAGAGCTATCCAACTTCGACAGAGTGT 907  
 Db 662 ACGTAAACGCTGCAATACCATCTTACCACCAAGAGCTATCCAACTTCGACAGAGTGT 603  
 QY 908 TTCAGAAATACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGA 967  
 Db 602 TTCAGAAATACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGA 543  
 QY 968 AAGGTGATGAGAAATGCTCAGCATCTGTAAGTGGCGCCCAAGCAACCAACAGCTT 1027  
 Db 542 AAGGTGATGAGAAATGCTCAGCATCTGTAAGTGGCGCCCAAGCAACCAACAGCTT 483  
 QY 1028 TCTTTGAGAGAACTTCATCAAGCCATGAAGGTTGTGAAGCTCGCCATGAAGCAATTGA 1087  
 Db 482 TCTTTGAGAGAACTTCATCAAGCCATGAAGGTTGTGAAGCTCGCCATGAAGCAATTGA 423  
 QY 1088 TCAGGATATGGTTCCCGTTCTGAATTCACATGATGATATCAAGCATTTCTATCAGA 1147  
 Db 422 TCAGGATATGGTTCCCGTTCTGAATTCACATGATGATATCAAGCATTTCTATCAGA 363  
 QY 1148 AGATGTATGGTATCTCCCTTTGCCAAGCCATCTGTGATGAACCAAGAGAGATGATGAGA 1207  
 Db 362 AGATGTATGGTATCTCCCTTTGCCAAGCCATCTGTGATGAACCAAGAGAGATGATGAGA 303  
 QY 1208 AAATCCTGGTGGCTCTTTGGAGAGAACTTAACATTCCTTGAATGCTCAAGAGCTATGA 1267  
 Db 302 AAATCCTGGTGGCTCTTTGGAGAGAACTTAACATTCCTTGAATGCTCAAGAGCTATGA 243  
 QY 1268 TCAGAGAGCTTTAATATATATATTTTCAATCTTAAAGCTTAAATAGGAAAGTTTCTTCAAC 1327  
 Db 242 TCAGAGAGCTTTAATATATATATTTTCAATCTTAAAGCTTAAATAGGAAAGTTTCTTCAAC 183  
 QY 1328 AGGATACAGTGTAGCTACCTPACATGCTGAAATAATATAGCTTTAATATATTTTATATT 1387  
 Db 182 AGGATACAGTGTAGCTACCTPACATGCTGAAATAATATAGCTTTAATATATTTTATATT 123  
 QY 1388 ATAACTCTGATATAGAGTAAGTCCATTTTAAATAAGTTTTCCTCCCAACCAATAAAA 1447  
 Db 122 ATAACTCTGATATAGAGTAAGTCCATTTTAAATAAGTTTTCCTCCCAACCAATAAAA 63  
 QY 1448 CCCTATACAACTTGTCTTAGTAAACATACATGAGAAAGAGTGTCTATGATGCTGAAATAA 1507  
 Db 62 CCCTATACAACTTGTCTTAGTAAACATACATGAGAAAGAGTGTCTAGGKRNGAAARNA 3  
 QY 1508 A 1508  
 Db 2 A 2

RESULT 9  
 AL541874  
 LOCUS  
 DEFINITION AL541874 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE007YN11

5-PRIME, mRNA sequence.  
 AL541874  
 AL541874.2 GI:30546467  
 EST.  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:12873366.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 5436.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DE007CG06QPl&cluster=5436.f. Contact :  
 Feng liang Email : fliang@lifetech.com URL : Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DE007CG06QPl.  
 Location/Qualifiers  
 1. 1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE007YN11"  
 /tissue\_type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."  
 Query Match 64.9%; Score 984; DB 9; Length 1201;  
 Best Local Similarity 99.0%; Pred. No. 3.4e-251;  
 Matches 990; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 QY 150 TCTCTTTAGTTCTTTGCAAGAGCTAGAGTAAGACACTTTTCAAAAATGCGCAATGGT 209  
 Db 62 TCTCTTTAGTTCTTTGCAAGAGCTAGAGTAAGACACTTTTCAAAAATGCGCAATGGT 121  
 QY 210 ATCAGAAATTCCTCAAGCAGGCTGTTTATTGAAATGAAGAGCAGGAATATGTTCAAAC 269  
 Db 122 ATCAGAAATTCCTCAAGCAGGCTGTTTATTGAAATGAAGAGCAGGAATATGTTCAAAC 181  
 QY 270 TGTGAAGTCATCAAGAGTGTCCGGATCAGCGGTGAGCCCTTCTCTACCTTCAATCC 329  
 Db 182 TGTGAAGTCATCAAGAGTGTCCGGATCAGCGGTGAGCCCTTCTCTACCTTCAATCC 241  
 QY 330 ATCCCTCGCATGCTGCTGCTTGCATAGGCCATTAATGTTAAAGGTGTGGATGAAGCAAC 389  
 Db 242 ATCCCTCGCATGCTGCTGCTTGCATAGGCCATTAATGTTAAAGGTGTGGATGAAGCAAC 301  
 QY 390 CATCATTTGACATTTCACTAAGCGAAACAATGCAAGCGTCAACAGATCAAGAGCAGATA 449  
 Db 302 CATCATTTGACATTTCACTAAGCGAAACAATGCAAGCGTCAACAGATCAAGAGCAGATA 361  
 QY 450 TCTCCAGAAACAGGAAAGCCCTGGATGAACACTGAAGAAAGCCCTTACAGGTCACCT 509  
 Db 362 TCTCCAGAAACAGGAAAGCCCTGGATGAACACTGAAGAAAGCCCTTACAGGTCACCT 421  
 QY 510 TGAGAGGTTGTTTGTGCTCTGCTTAAACAACTCCAGCGCAATTTGATGCTGATGAACCTCG 569  
 Db 422 TGAGAGGTTGTTTGTGCTCTGCTTAAACAACTCCAGCGCAATTTGATGCTGATGAACCTCG 481  
 QY 570 TGCTCCCATGAAGGCGCTTGGAACTGATGAAGATCTCTAATTGAGATTTTGGCATCAAG 629

FEATURES  
 source

ORIGIN

482	DB	TTGCTGCCATGAAGGGCTTCTGAACTGATGAAGATACCTCTAATTGAGATTTTGGCATCAAG	541
630	QY	AACCTAACAAAGAATAACAGACACATTAACAGGGTCTACAGAGAGAACTGAAGAGAGATCT	689
542	DB	AACCTAACAAAGAATAACAGACATTAACAGGGTCTACAGAGAGAACTGAAGAGAGATCT	601
690	QY	GGCCAAAGACATAAACCCTCAGACACATCTGAGATTTTCGAAACGCTTTGCTTCTCTTGC	749
602	DB	GGCCAAAGACATAAACCCTCAGACACATCTGAGATTTTCGAAACGCTTTGCTTCTCTTGC	661
750	QY	TAAAGGTGACCGATCTGAGGACATTTGGTGATGAAGACATTCGCTGATTCAGATGCCAG	809
662	DB	TAAAGGTGACCGATCTGAGGACATTTGGTGATGAAGACATTCGCTGATTCAGATGCCAG	721
810	QY	GGCCTTTGTATGAAGCAGGAGAAAGGAGGACAGACGTAAACGTGTTCATACCAT	869
722	DB	GGCCTTTGTATGAAGCAGGAGAAAGGAGGACAGACGTAAACGTGTTCATACCAT	781
870	QY	CCATTACCACAGAGACTATCCACAACTTCGACAGTGTTCAGAAATACACCAAGTACAG	929
782	DB	CCATTACCACAGAGACTATCCACAACTTCGACAGTGTTCAGAAATACACCAAGTACAG	841
930	QY	TAAAGCATGACATGAACAAAGTTCTCGACCTCGAGCTTGAAGAGTGACATTCAGAAATCGCT	989
842	DB	TAAAGCATGACATGAACAAAGTTCTCGACCTCGAGCTTGAAGAGTGACATTCAGAAATCGCT	901
990	QY	CACAGCTATCGTGAAGTGGCCACAAGCAACCAAGCTTCTTTTCAGAGAGAGCTTCATCA	1049
902	DB	CACAGCTATCGTGAAGTGGCCACAAGCAACCAAGCTTCTTTTCAGAGAGAGCTTCATCA	961
1050	QY	AGCCATCAAAAGGTCTCGAACTCGCCCATMAGGCATTCATCAGGATTTATGGTTCCCGCTTC	1109
962	DB	AGCCATCAAAAGGTCTCGAACTCGCCCATMAGGCATTCATCAGGATTTATGGTTCCCGCTTC	1021
1110	QY	TGAAATTTGACATGAATGATATCAAAAGCATTTCTATCAGAAG	1149
1022	DB	TGAAATTTGACATGAATGATATCAAAAGCATTTCTATCAGAATG	1061

RESULT 10	AL574991/c	AL574991	1136 bp	mrna	linear	EST 01-JUN-2003
LOCUS	AL574991	AL574991	1136 bp	mrna	linear	EST 01-JUN-2003
DEFINITION	AL574991	AL574991	1136 bp	mrna	linear	EST 01-JUN-2003
ACCESSION	AL574991	AL574991	1136 bp	mrna	linear	EST 01-JUN-2003
VERSION	AL574991.2	AL574991.2	1136 bp	mrna	linear	EST 01-JUN-2003
KEYWORDS	EST.	EST.	1136 bp	mrna	linear	EST 01-JUN-2003
SOURCE	Homo sapiens (human)	Homo sapiens (human)	1136 bp	mrna	linear	EST 01-JUN-2003
ORGANISM	Homo sapiens	Homo sapiens	1136 bp	mrna	linear	EST 01-JUN-2003
REFERENCE	1 (bases 1 to 1136)	1 (bases 1 to 1136)	1136 bp	mrna	linear	EST 01-JUN-2003
AUTHORS	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.	1136 bp	mrna	linear	EST 01-JUN-2003
TITLE	Full-length cDNA libraries and normalization	Full-length cDNA libraries and normalization	1136 bp	mrna	linear	EST 01-JUN-2003
JOURNAL	Unpublished (2001)	Unpublished (2001)	1136 bp	mrna	linear	EST 01-JUN-2003
COMMENT	On Feb 16, 2001 this sequence version replaced gi:12935721.	On Feb 16, 2001 this sequence version replaced gi:12935721.	1136 bp	mrna	linear	EST 01-JUN-2003

```

/db_xref="taxon:9606"
/clone="CS0D10649E07"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match          64.1%; Score 971.4; DB 9; Length 1136;
Best Local Similarity 93.8%; Pred. No. 7.7e-248;
Matches 1027; Conservative 15; Mismatches 46; Indels 7; Gaps 3;

QY 396 TGACATCTTAACCTAAGCGAAACAATGCACAGCGTCTCAACAGATCAACAGACGACATATCTCCA 455
DB 1088 TGAGCCACCTATTACATYTTAAAAAATAAAWMAAGGYAACARACAAASMGCAATTCYC 1029

QY 456 GGAACAGGAAAGCCCTGGATGAACACTGAGAAAGGCCCTTACAGGTACACCTGAGGA 515
DB 1028 MGGAAACAGAAAGCCCTGTATRAACAKCAAGAAR-----SCCTACRCGTCACTTTGAGGA 974

QY 516 GGTGTGTTTAGCTCTGTCTAAAACTCCAGCGCAATTGTGCTGTGAGAACTTCGTGCTGC 575
DB 973 -GTTGTTTTAGCTCTGTCTAAAACTCCAGCGCAATTGTGCTGTGAGAACTTCGTGCTGC 916

QY 576 CATGAAGGGCCCTTGGAACTGATGAAGATCTCTAATTGAGATTTTGGCATCAAGAACTAA 635
DB 915 CATGAAGGGCCCTTGGAACTGATGAAGATCTCTAATTGAGATTTTGGCATCAAGAACTAA 856

QY 636 CAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGGAACTGAAGAGAGATCTGGCCAA 695
DB 855 CAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGGAACTGAAGAGAGATCTGGCCAA 796

QY 696 AGACATAACCTCAGACACATCTGGAGATTTTCGAGAGCCTTCGTTCTCTTGTCTAAGGG 755
DB 795 AGACATAACCTCAGACACATCTGGAGATTTTCGAGAGCCTTCGTTCTCTTGTCTAAGGG 736

QY 756 TGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTGATTCAGATGCCAGGGCCTT 815
DB 735 TGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTGATTCAGATGCCAGGGCCTT 676

QY 816 GTATGAAGCAGGAGAAAGGAGAAAGGGGACAGAGCTGAACGTTCAATACCATCCTTAC 875
DB 675 GTATGAAGCAGGAGAAAGGAGAAAGGGGACAGAGCTGAACGTTCAATACCATCCTTAC 616

QY 876 CACCAGAAGCTATCCCAACCTTCGAGAGTGTTCAGAAATACACCAAGTACAGTAAGCA 935
DB 615 CACCAGAAGCTATCCCAACCTTCGAGAGTGTTCAGAAATACACCAAGTACAGTAAGCA 556

QY 936 TGACATGAAACAAAGTCTCGACCTGGAGTTGAAGGTGACATTCAGAGAAATGCGCTCACAGC 995
DB 555 TGACATGAAACAAAGTCTCGACCTGGAGTTGAAGGTGACATTCAGAGAAATGCGCTCACAGC 496

QY 996 TATCGTGAAGTGGGCCCAAGCAAAACAGCTTTCTTTTCGAGAGAGCTTCATCAAGCCAT 1055
DB 495 TATCGTGAAGTGGGCCCAAGCAAAACAGCTTTCTTTTCGAGAGAGCTTCATCAAGCCAT 436

QY 1056 GAAAGGTGTTGGAACTCGCCATAGGCATTTGATCAGGATTTATGTTTCCCGTTCTGAAAT 1115
DB 435 GAAAGGTGTTGGAACTCGCCATAGGCATTTGATCAGGATTTATGTTTCCCGTTCTGAAAT 376

QY 1116 TGACATGAATGATATCAAGCATTTCTATCAGAGATGTATGGTATCTCCCTTTGCCAAGC 1175
DB 375 TGACATGAATGATATCAAGCATTTCTATCAGAGATGTATGGTATCTCCCTTTGCCAAGC 316

QY 1176 CATCCTGTGTAACCAAGAGAGATGAGAAATCCTGGTGGCTCTTTTGTGGAGGAAA 1235
DB 315 CWTCTGTGATGAACCAAGAGAGATTTATGAGAAATCCTGGTGGCTCTTTTGTGGAGGAAA 256

QY 1236 CTAACACATCCCTTGTGGTCTCAAGCTATGATCAGAGACTTTAATATATATTTTCAT 1295
DB 255 CTAACACATCCCTTGTGGTCTCAAGCTATGATCAGAGACTTTAATATATATTTTCAT 196

```

```

QY 1296 CCTATAAGCTTAAATAGGAAGTTCTTCAACAGGATTACAGTGTAGCTACCTACATGCT 1355
Db 195 CCTATAGCTTAAATAGGAAGTTCTTCAACAGGATTACAGTGTAGCTACCTACATGCT 136
QY 1356 GAAATATATAGCTTAAATCAATCTTTTATATATATACTCTGTATATAAGAGATAAGTCCA 1415
Db 135 GAAATATATAGCTTAAATCAATCTTTTATATATATACTCTGTATATAAGAGATAAGTCCA 1415
QY 1416 TTTTATAAATAGTTTCCCAAAACCAATAAACCCTATACAGTTGTTCTAGTAAACAATA 1475
Db 75 TTTTATAAATAGTTTCCCAAAACCAATAAACCCTATACAGTTGTTCTAGTAAACAATA 1475
QY 1476 CATGGAAGATGTC 1490
Db 15 CATGGAAGATGTC 1

RESULT 11
AL553129
LOCUS
DEFINITION
AL553129 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CSOD1074116 5-PRIME, mRNA sequence.
ACCESSION
AL553129
VERSION
AL553129.2 GI:31274943
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12892678.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5436.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOD1074BE080P1&cluster=5436.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSOD1074BE080P1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1074Y116"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-ciligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 63.3%; Score 959.8; DB 9; Length 1201;
Best Local Similarity 95.5%; Pred. No. 9.7e-245;
Matches 101; Conservative 8; Mismatches 35; Indels 5; Gaps 3;

QY 142 GAAGATTTCTTTAGTTCTTTGCAAGAGGTAGAGATAAGACATTTTCAAAATG 201
Db 66 GAAGATTTCTTTAGTTCTTTGCAAGAGGTAGAGATAAGACATTTTCAAAATG 125
QY 202 GCAATGGTATCAGATTCTCTAAGCAGCGCTGTTTATTGAAATGAAGACAGGATAT 261
Db 126 GCAATGGTATCAGATTCTCTAAGCAGCGCTGTTTATTGAAATGAAGACAGGATAT 195

```

```

QY 262 GTTCAAACTGTGAAGTTCATCCAAAGTGG-TCCCGGATCAGGGGTGAGCCCTATCTCTAC 320
Db 186 GTTCAAACTGTGAAGTTCATCCAAAGTGGATCCCGGATCAGGGGTGAGCCCTATCTCTAC 245
QY 321 CTTCAATCCATCTCTCGGATGTCTGCTTGGCATTGCAATAAGGCCATTAATGGTTAAAGGTGGA 380
Db 246 CTTCAATCCATCTCTCGGATGTCTGCTTGGCATTGCAATAAGGCCATTAATGGTTAAAGGTGGA 305
QY 381 TGAAGCAACCATCATTTGACATTTCTAAGCAACCAATGACACAGCGTCAACAGATCAA 440
Db 306 TGAAGCAACCATCATTTGACATTTCTAAGCAACCAATGACACAGCGTCAACAGATCAA 365
QY 441 AGCAGCATATCTCCAGAAAACAGAAAGCCCTTGGATGAAACACTGAAAGAGCCCTTAC 500
Db 366 AGCAGCATATCTCCAGAAAACAGAAAGCCCTTGGATGAAACACTGAAAGAGCCCTTAC 425
QY 501 AGGTCACTTTGAGGAGGTTGTTTTAGCTCTGTCTTAAAAAACCTCCAGCGCAATTTGATCTGA 560
Db 426 AGGTCACTTTGAGGAGGTTGTTTTAGCTCTGTCTTAAAAAACCTCCAGCGCAATTTGATCTGA 485
QY 561 TGAATTCGTGCTGCCATGAGGGCCCTTGGAACTGATGAAGTACTCTAATTTGAGATTTT 520
Db 486 TGAATTCGTGCTGCCATGAGGGCCCTTGGAACTGATGAAGTACTCTAATTTGAGATTTT 545
QY 621 GGCATCAAGAACTAAACAAAGAAATCGAGACATTAACAGGGTCTACAGAGAGAACTGAA 680
Db 546 GGCATCAAGAACTAAACAAAGAAATCGAGACATTAACAGGGTCTACAGAGAGAACTGAA 605
QY 681 GAGAGATCTGGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAAGCGTTGCT 740
Db 606 GAGAGATCTGGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAAGCGTTGCT 665
QY 741 TTCTCTTTGTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTGGCTGATT 800
Db 666 TTCTCTTTGTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTGGCTGATT 725
QY 801 AGATGCCAGGGCTTTGATGAGCAGGAGAAAGGAGAAAGGGGACAGAGCTAAACGTGT 860
Db 726 AGATGCCAGGGCTTTGATGAGCAGGAGAAAGGAGAAAGGGGACAGAGCTAAACGTGT 785
QY 861 CAATACCATTCTTACCACCAAGAGCTATCCAACTTCGACAGAGTCTTTTCAGAAATACAC 920
Db 786 CAATACCATTCTTACCACCAAGAGCTATCCAACTTCGACAGAGTCTTTTCAGAAATACAC 845
QY 921 CAATCAGTAGTAAAGTACATGACATGAACAAAGTTCTGGACCTGGAGTTCAGAGGTGACATTC 980
Db 846 CAATCAGTAGTAAAGTACATGACATGAACAAAGTTCTGGACCTGGAGTTCAGAGGTGACATTC 905
QY 981 GAAATGCTCTCAGCTATCGTGAAGTGGCCCAACAGCAAAACCAAGCTTTCTTTGCAGAGAA 1040
Db 906 GAAATGCTCTCAGCTATCGTGAAGTGGCCCAACAGCAAAACCAAGCTTTCTTTGCAGAGAA 965
QY 1041 GCTTCATCAAGCCATGAAAGGTGTGGAACTGCCATAGGCAATTCATCAGGATTTATGTT 1100
Db 966 GCTTCATCAAGCCATGAAAGGTGTGGAACTGCCATAGGCAATTCATCAGGATTTATGTT 1025
QY 1101 TTCCCGTTCTGAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1160
Db 1026 TTCCCGTTCTGAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081
QY 1161 CTCCTTTGCGCAAGCCATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
Db 1082 CTCCTTTGCGCAAGCCATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1120

```

```

RESULT 12
AL544959
LOCUS
DEFINITION
AL544959 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CSOD1012P05 5-PRIME, mRNA sequence.
ACCESSION
AL544959
VERSION
AL544959.2 GI:31266800
KEYWORDS
EST.

```



primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 63.1%; Score 956; DB 9; Length 1164;  
 Best Local Similarity 94.2%; Pred. No. 9.9e-244;  
 Matches 1044; Conservative 13; Mismatches 38; Indels 13; Gaps 6;

QY 130 GAAATCTTCAGAGAGAAATTTCTTTAGTTAGTTCTTTGAGAGAGTGAAGATGAAGACACT 189  
 DB 49 GAAATCCCGGATAGAAATTTCTTTAGTTAGTTCTTTGAGAGAGTGAAGATGAAGACACT 108

QY 190 TTTTCAAAATGGCAATCGTATCAGAAATCTCTCAAGCAGGCTGCTTTATTGAAATGAA 249  
 DB 109 TTTTCAAAATGGCAATCGTATCAGAAATCTCTCAAGCAGGCTGCTTTATTGAAATGAA 168

QY 250 GAGCAGGAATATGTTCAAACTGTGAAGTCAATCAAGAGTGTCCGATCAGCGGTGAGC 309  
 DB 169 GAGCAGGAATATGTTCAAACTGTGAAGTCAATCAAGAGTGTCCGATCAGCGGTGAGC 228

QY 310 CCTATCTCTACCTTCAATCCATCTCTCGATGTCGTGCTTGCATTAAGGCCATTAATGTT 369  
 DB 229 CCTATCTCTACCTTCAATCCATCTCTCGATGTCGTGCTTGCATTAAGGCCATTAATGTT 288

QY 370 AAAGTGTGGATGAAGCAACCATCATTCACATTTCACTAAGCGAAACAAATGCACAGCT 429  
 DB 289 AAAGTGTGGATGAAGCAACCATCATTCACATTTCACTAAGCGAAACAAATGCACAGCT 348

QY 430 CNAACAGATCAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACACTCAAG 489  
 DB 349 CNAACAGATCAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACACTCAAG 408

QY 490 AAAGCCCTTACAGGTCACTCTGAGGAGGTGTTGTTAGCTTCTGTAATAAATCCAGCGCAA 549  
 DB 409 AAAGCCCTTACAGGTCACTCTGAGGAGGTGTTGTTAGCTTCTGTAATAAATCCAGCGCAA 468

QY 550 TTTGATGCTGATGAATCTGCTGCTCCATGAAGGCTTGAACCTGATGAAGTACTCTA 609  
 DB 469 TTTGATGCTGATGAATCTGCTGCTCCATGAAGGCTTGAACCTGATGAAGTACTCTA 528

QY 610 ATTGAGATTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGGGTCTACAGA 669  
 DB 529 ATTGAGATTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGGGTCTACAGA 588

QY 670 GAGGAATCTGAAGAGATCTGCGCAAGACATPAACCTCAGACACATCTGGAGTTTCGG 729  
 DB 589 GAGGAATCTGAAGAGATCTGCGCAAGACATPAACCTCAGACACATCTGGAGTTTCGG 648

QY 730 AACGCTTTCTTCTTCTTCTAAGGCTGACCATCTGAGGACTTTGGTGTGAATGAAGAC 789  
 DB 649 AACGCTTTCTTCTTCTTCTAAGGCTGACCATCTGAGGACTTTGGTGTGAATGAAGAC 708

QY 790 TTGGCTGATTCAGATGCCAGGCTTGTATGAAGCAGAGAAAGAGAAAGGGGACAGAC 849  
 DB 709 TTGGCTGATTCAGATGCCAGGCTTGTATGAAGCAGAGAAAGAGAAAGGGGACAGAC 768

QY 850 GTAACCGTCTCAATACCATCTTACCACAGAGCTATCCAACTCTCGAGAGTGTGT 909  
 DB 769 GTAACCGTCTCAATACCATCTTACCACAGAGCTATCCAACTCTCGAGAGTGTGT 828

QY 910 CAGAAATACACCAAGTACAGTAAAGCATGACATGAACAAAGTTCTGAGCTGGAGTTGAAA 969  
 DB 829 CAGAAATACACCAAGTACAGTAAAGCATGACATGAACAAAGTTCTGAGCTGGAGTTGAAA 888

QY 970 GGTGACATTTGAGAAATGCTTACAGCTATCTGTAAGTGGCCACAGCAAAACCACTTTC 1029  
 DB 889 GGTGACATTTGAGAAATGCTTACAGCTATCTGTAAGTGGCCACAGCAAAACCACTTTC 948

QY 1030 TTTTCAGAGAGAGCTTTCATCAGCCATGAAGGTTGTTGAACTCGCCATTAAGCATTTGATC 1089  
 DB 949 TTTTCAGAGAGAGCTTTCATCAGCCATGAAGGTTGTTGAACTCGCCATTAAGCATTTGATC 1007

QY 1090 AGGATTATGTTTCCGTTCTGAATTCGACATGATGATCAAAAGCATTTCTATCAGAAG 1149  
 DB 1008 AGATATATGG-TTCGGTCTTGAATTCGAAGGA---TGAAATTAAGMATCTATWAG-AR 1062

QY 1150 ATGTATGATATCTCCCTTTGCCAAGCCATCTCGATGAACCAAGAGGAGATGAGAAA 1209  
 DB 1063 ATGTATGGAATTCCTTT---GCMASCATCTCGGATGA---ACAAGGGGTTTGTGAAA 1115

QY 1210 ATCTCTGGTGGCTCTTTGTGGAGAAACT 1237  
 DB 1116 ATTTGGKGGTTTTTGGGAAT 1143

RESULT 14  
 AL576223/c  
 LOCUS  
 DEFINITION  
 AL576223 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0D1074Y116 3-PRIME, mRNA sequence.  
 ACCESSION  
 AL576223  
 VERSION  
 AL576223.2 GI:31314513  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1. (bases 1 to 1201)  
 Li, W.B., Gruber, C., Jessee, J. and Pollayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 16, 2001 this sequence version replaced gi:12938153.  
 COMMENT  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by Life technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 5436.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0D1074B08NP1&cluster=5436.f. Contact :  
 Feng Liang Email: fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0D1074B08NP1.

FEATURES  
 Location/Qualifiers  
 1..1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1074Y116"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 63.0%; Score 955.8; DB 9; Length 1201;  
 Best Local Similarity 97.4%; Pred. No. 1.1e-243;  
 Matches 994; Conservative 8; Mismatches 14; Indels 5; Gaps 3;

QY 470 CCTGTGATGAACACATGAAGAACCCCTTACAGGTCACTTGTGAGGAGTTGTTTGTAGCTC 529  
 DB 1016 CCTGTGTGAACATGTGAAGAACCC--TTACAGDCACCTWRAGGAGTTGTTAA--CT 961

QY 530 TGTCTAAAACTCCAGGCAATTTGATGCTGATGAACCTTGTGCTGCATGAAGGCTTTC 589  
 DB 960 CTVTAAAACTCCAGGCAATTTGATGCTGATGAAC--TCGTGCTGCATGAAGGCTTTC 902

QY 590 GAATGATGAAGATCTCTTAATTGAGATTTTGCATCAAGAACTAACAAGAAATCAGAG 649  
 DB 901 GAATGATGAAGATCTCTMAAARAGATTTTGCATCAAGAACTAACAAGAAATCAGAG 842

QY 650 ACATTAACAGGCTCTACAGAGAGGAACTGAAGAGAGATCTGGCCAAAGACATTAACCTCAG 709







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model  
Run on: March 10, 2004, 08:38:17 ; Search time 6354.47 Seconds  
(without alignments)  
17011.249 Million cell updates/sec

Title: US-10-084-817-18  
Perfect score: 2494  
Sequence: 1 Gcgtgaggatcgcgcgtt.....ccattaaaaa 2494

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	2477.8	99.4	2488	9	BC000013		BC000013 Homo sapi
2	2457.4	98.5	2549	6	I09499		I09499 Sequence 20
3	2456	98.5	2496	6	AX281786		AX281786 Sequence
4	2446.2	98.1	2499	9	HGIGFBP3M		X64875 H. sapiens m
5	2446.2	98.1	2550	6	AR364291		AR364291 Sequence
6	2445.8	98.1	2497	9	BC018962		BC018962 Homo sapi
7	2430.4	97.4	2474	9	HUMGFBPA		M31159 Human growt
8	2361.8	94.7	2448	6	AX713809		AX713809 Sequence
9	2361.8	94.7	2448	6	AK055632		AK055632 Homo sapi
10	2309.4	92.6	2430	9	AK122873		AK122873 Homo sapi
11	2283.4	91.6	2412	9	AK075483		AK075483 Homo sapi
12	2248.6	90.2	2364	9	AK095907		AK095907 Homo sapi
13	2240.8	89.8	2343	9	AK095869		AK095869 Homo sapi
14	2229.4	89.4	2386	9	AK095408		AK095408 Homo sapi
15	2209.8	88.6	2372	9	AK058274		AK058274 Homo sapi
16	2162.6	86.7	2355	6	AX833836		AX833836 Sequence
17	2162.6	86.7	2355	9	AK095900		AK095900 Homo sapi
18	2156.8	86.5	2312	9	AK091745		AK091745 Homo sapi
19	2152.6	86.3	2343	9	AK095991		AK095991 Homo sapi
20	2143.4	85.9	2191	6	AX014094		AX014094 Sequence
21	2143.4	85.9	2191	6	BD204664		BD204664 Human nuc
22	2101.4	84.3	2327	9	AK127537		AK127537 Homo sapi
23	2099.8	84.2	2327	6	AX747551		AX747551 Sequence
24	2099.8	84.2	2327	9	AK092447		AK092447 Homo sapi
25	2055.4	82.4	2208	9	AK125224		AK125224 Homo sapi
26	2050.8	82.2	2207	9	AK097201		AK097201 Homo sapi
27	2024.8	81.2	2242	9	AK091706		AK091706 Homo sapi
28	1977.6	79.3	2192	6	AX833876		AX833876 Sequence
29	1977.6	79.3	2192	9	AK095974		AK095974 Homo sapi
30	1968	78.9	2141	9	AK123735		AK123735 Homo sapi
31	1943.8	77.9	2173	6	AX833820		AX833820 Sequence
32	1943.8	77.9	2173	9	AK095864		AK095864 Homo sapi
33	1898	76.1	2099	9	AK129951		AK129951 Homo sapi
34	1882.4	75.5	2011	9	AK130884		AK130884 Homo sapi
35	1474.6	59.1	3708	9	HSM807112		EX640961 Homo sapi
36	1471.8	59.0	69887	9	AC091524		AC091524 Homo sapi
37	1460.6	58.6	14539	9	AC146152		AC146152 Pan trogl
38	1460.6	58.6	182079	2	AC146117		AC146117 Pan trogl
39	1459.8	58.5	10884	6	AX409699		AX409699 Sequence
40	1459.8	58.5	10884	9	HUMIBP3		M35878 Homo sapien
41	1426.4	57.2	2438	11	G06425		G06425 human STS W
42	1104.6	44.3	2430	4	AF305199		AF305199 Bos tauru
43	968.4	38.8	2432	10	BC058261		BC058261 Mus muscu
44	914.4	36.7	916	6	AR272361		AR272361 Sequence
45	914.4	36.7	916	6	AR275942		AR275942 Sequence

ALIGNMENTS

RESULT 1  
BC000013  
LOCUS  
DEFINITION  
BC000013 Homo sapiens insulin-like growth factor binding protein 3, mRNA  
(cdna clone MGC:2305 IMAGE:350666), complete cds.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2488)  
REFERENCE  
AUTHORS  
Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altechul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,K.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E., Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 2488)

Strausberg,R.

Direct Submission

Submitted (03-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Aug 20, 2003 this sequence version replaced gi:12652546.

Contact: MGC help desk

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: [amadan@systemsbiology.org](mailto:amadan@systemsbiology.org)

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 7 Row: e Column: 4

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19923110.

Location/Qualifiers

1..2488

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC12305 IMAGE:3506666"

/tissue\_type="Kidney, renal cell adenocarcinoma"

/clone\_lib="NIH MGC\_14"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

1..2488

/gene="IGFBP3"

/note="Synonym: IBP3"

/db\_xref="LocusID:3486"

/db\_xref="MIM:146732"

99..974

/codon\_start=1

/product="insulin-like growth factor binding protein 3"

/protein\_id="AAH00013.1"

/db\_xref="GI:12652547"

/db\_xref="LocusID:3486"

/translation="MQRARPTLWAALTLVLRLGPPVARAGASGGGLGPPVRCEPCDAPALAQCPAPVACALVREPCGCLTALSGEQPCGIYTERCGSLRCOPSPDEARPLQALLDGRGLCNVALSVRLRAYLLPAPAPGNASESEDSRAGSVSPSVSTHVRSDPKPHLSKTIILIKGHAKDSORVKVDYESQSTDTONFSSBSKRETEYGPCKREMEPTLNHLKFLNVLSPSGVHPNCDKKGFKYKKQCRPSKRGKFCWCVDYKQPLPGYT

TKGKEDVHCYSMQSK"

misc\_feature 210..446

/note="IB; Region: Insulin growth factor-binding protein homologues"

/db\_xref="CDD:smart00121"

735..953

/note="Thyroglobulin\_1; Region: Thyroglobulin type-1 repeat. Thyroglobulin type 1 repeats are thought to be involved in the control of proteolytic degradation. The domain usually contains six conserved cysteines. These form three disulphide bridges. Cysteines 1 pairs with 2, 3 with 4 and 5 with 6"

/db\_xref="CDD:pfam00086"

ORIGIN

Query Match 99.4%; Score 2477.8; DB 9; Length 2488;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2479; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 GCGCTTCTCGCTGGATTCCACAGCTTGGCCCGGTGACTGTCCGCCATCCCTGGCGG 73

DB 1 GCGCTTCTCGCTGGATTCCACAGCTTGGCCCGGTGACTGTCCGCCATCCCTGGCGG 60

QY 74 CCCAGCTCCCAAGCAGCTGCCCGGTTCAGGGTTCATGCGAGGGGGCGACCCAGCG 133

DB 61 CCCAGCTCCCAAGCAGCTGCCCGGTTCAGGGTTCATGCGAGGGGGCGACCCAGCG 120

QY 134 TCTGGGCCCGCTCGGTGACTCTGCTGGTGTCTCTCGCGGGCGCGCGGTGGCGGGCTG 193

DB 121 TCTGGGCCCGCTCGGTGACTCTGCTGGTGTCTCTCGCGGGCGCGCGGTGGCGGGCTG 180

QY 194 GCGCAGCTCGGGGGGGCTTGGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGT 253

DB 181 GCGCAGCTCGGGGGGGCTTGGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGT 240

QY 254 TGGCCAGTGGCGCGCTCCGCCCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGG 313

DB 241 TGGCCAGTGGCGCGCTCCGCCCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGG 300

QY 314 GTGTGCTGCTGAGTGGCGACTGAGGAGGGCGAGCGGTTCGGTTCGGTTCGGTTCGGTTCGG 373

DB 301 GTGTGCTGCTGAGTGGCGACTGAGGAGGGCGAGCGGTTCGGTTCGGTTCGGTTCGGTTCGG 360

QY 374 GTGTGCTGCTGAGTGGCGACTGAGGAGGGCGAGCGGTTCGGTTCGGTTCGGTTCGGTTCGG 433

DB 361 GTGTGCTGCTGAGTGGCGACTGAGGAGGGCGAGCGGTTCGGTTCGGTTCGGTTCGGTTCGG 420

QY 434 TGGACGGCGCGGGCTCTCGCTCAACGCTAGTGGCGCTGAGCGCGCTGCGCGCTTACCTGC 493

DB 421 TGGACGGCGCGGGCTCTCGCTCAACGCTAGTGGCGCTGAGCGCGCTTACCTGC 480

QY 494 TGGACGGCGCGGGCTCTCGCTCAACGCTAGTGGCGCTGAGCGCGCTTACCTGC 553

DB 481 TGGACGGCGCGGGCTCTCGCTCAACGCTAGTGGCGCTGAGCGCGCTTACCTGC 540

QY 554 GTGTGAGAGCGCGCTCGCTCTCCAGCAGCGCGGTGCTGATCCCAAGTTCACCCCG 613

DB 541 GTGTGAGAGCGCGCTCGCTCTCCAGCAGCGCGGTGCTGATCCCAAGTTCACCCCG 600

QY 614 TCCATTCAAGATATCATCATCAAGAGGGCGTCTAAAGCAGCGCGCTACCAAG 673

DB 601 TCCATTCAAGATATCATCATCAAGAGGGCGTCTAAAGCAGCGCGCTACCAAG 660

QY 674 TTGACTACGAGTCTCAGACAGATACCCAGACTTCTCTCCGAGTCCAGCGGGAGA 733

DB 661 TTGACTACGAGTCTCAGACAGATACCCAGACTTCTCTCCGAGTCCAGCGGGAGA 720

QY 734 CAGAATATGTCCTCGCTCGCTAGGAATGGAGACACACTGAATCACCCTGAATTCCTCA 793

DB 721 CAGAATATGTCCTCGCTCGCTAGGAATGGAGACACACTGAATCACCCTGAATTCCTCA 780

QY 794 ATGTGCTGAGTCCCGGGGTGTACACATTTCCCACTGTGCAAGAGGGATTTTAAAGA 853

DB 781 ATGTGCTGAGTCCCGGGGTGTACACATTTCCCACTGTGCAAGAGGGATTTTAAAGA 840

854 AAAAGCAGTGTGCGCCCTTCCAAAGCGCAGGAGCGGGGCTTCTCTCGTGTGTGCGATAAGT 913  
841 AAAAGCAGTGTGCGCCCTTCCAAAGCGCAGGAGCGGGGCTTCTCTCGTGTGTGCGATAAGT 900  
914 ATGGGCAGCCTCTCCAGGCTACACCAACCAAGGAGGAGGAGCGTGCATCTGCTACAGCA 973  
901 ATGGGCAGCCTCTCCAGGCTACACCAACCAAGGAGGAGGAGCGTGCATCTGCTACAGCA 960  
974 TGCAGAGCAATAGACGCCCTCGCGCAAGGTTAATGTGGAGCTCAAAATATGCCCTATTATTG 1033  
961 TGCAGAGCAATAGACGCCCTCGCGCAAGGTTAATGTGGAGCTCAAAATATGCCCTATTATTG 1020  
1034 CACAAAGAGCTGCCAAGGACATGACCAAGGAGCTGGCTACAGCTCGATTTATATTTCTGT 1093  
1021 CACAAAGAGCTGCCAAGGACATGACCAAGGAGCTGGCTACAGCTCGATTTATATTTCTGT 1080  
1094 TTGTGGTGAATGATTTTTTTTAAACCAAGTTTAGAAAGAGGTTTTTGAATGCCCTAT 1153  
1081 TTGTGGTGAATGATTTTTTTTAAACCAAGTTTAGAAAGAGGTTTTTGAATGCCCTAT 1140  
1154 GGTTCTTTTGAATGGTAACTTGAGCATCTTTTCACTTTCAGTACGACGAAGAGCAG 1213  
1141 GGTTCTTTTGAATGGTAACTTGAGCATCTTTTCACTTTCAGTACGACGAAGAGCAG 1200  
1214 TTTGAATTTTCTTGTGCGCTTCCCTATCAAAATATTTCAGAGACTCGAGCAGCACCCAGAC 1273  
1201 TTTGAATTTTCTTGTGCGCTTCCCTATCAAAATATTTCAGAGACTCGAGCAGCACCCATAC 1260  
1274 TTCATGCGCGGTGAATGCTCACCACATGTTGGTGAAGCGGCCAGCACTGACTTTGT 1333  
1261 TTCATGCGCGGTGAATGCTCACCACATGTTGGTGAAGCGGCCAGCACTGACTTTGT 1320  
1334 GACTTAGCGCGCTGTTGTGCTATGTAGAAACACGCTTCACCCCACTCCCGTACAGT 1393  
1321 GACTTAGCGCGCTGTTGTGCTATGTAGAAACACGCTTCACCCCACTCCCGTACAGT 1380  
1394 GCGCAGGCTTTATCGAAGATAGAAAACCTTTAAACCCCGTCTACCGAGCATCCCAA 1453  
1381 GCGCAGGCTTTATCGAAGATAGAAAACCTTTAAACCCCGTCTACCGAGCATCCCAA 1440  
1454 CGCATGCTCTGAGCTCACAGCCTTCTGTGTGTCTATTTCTGAAACAGGGGCTGGATC 1513  
1441 CGCATGCTCTGAGCTCACAGCCTTCTGTGTGTCTATTTCTGAAACAGGGGCTGGATC 1500  
1514 CCTCAACCAAGAGAAATGTTTATGCTTCAAGTGAACCTGCTGCTTGGGACATATTGA 1573  
1501 CCTCAACCAAGAGAAATGTTTATGCTTCAAGTGAACCTGCTGCTTGGGACATATTGA 1560  
1574 GAAATTAAGTGGAGTCTTCTTGTAAATAATATGATCTAAGATGTTCTAGGGCAC 1633  
1561 GAAATTAAGTGGAGTCTTCTTGTAAATAATATGATCTAAGATGTTCTAGGGCAC 1620  
1634 TCTGGGAACCTATAAGGCGAGTTATTTGGGCGCTCTCTTTCAGGAATCTTCTGAGAC 1693  
1621 TCTGGGAACCTATAAGGCGAGTTATTTGGGCGCTCTCTTTCAGGAATCTTCTGAGAC 1680  
1594 ATGGCCAGTTCGAGGCGCAGGATGGCTTTTGTGGCGCCGCTGGGTAGGGAGCAG 1753  
1681 ATGGCCAGTTCGAGGCGCAGGATGGCTTTTGTGGCGCCGCTGGGTAGGGAGCAG 1740  
1754 AGACAGGAGAGTTCAGCTCCACATTCAGAGCATCAAGTAATGGCACAATTTCTTC 1813  
1741 AGACAGGAGAGTTCAGCTCCACATTCAGAGCATCAAGTAATGGCACAATTTCTTC 1800  
1814 GGATGACTCGAAGAAATAGTGTTTTGTAGTTCAACACTCAAGCAGGTTATTTCTGA 1873  
1801 GGATGACTCGAAGAAATAGTGTTTTGTAGTTCAACACTCAAGCAGGTTATTTCTGA 1860  
1874 GGATAAGCTCTTTAAAGGCAAGCTTTATTTTTCATCTCTCATCTTTTGTCTCTTAGCA 1933  
1861 GGATAAGCTCTTTAAAGGCAAGCTTTATTTTTCATCTCTCATCTTTTGTCTCTTAGCA 1920

1934 CAATGTAAAAAGAAATAGTAATATCAGAAACAGGAAGGAATGGCTTGTCTGGGAGGCC 1993  
1921 CAATGTAAAAAGAAATAGTAATATCAGAAACAGGAAGGAATGGCTTGTCTGGGAGGCC 1980  
1994 ATCCAGGACACTGGGAGCACATAGAGATTACCCATGTTTGTGAACCTTAGAGTCATTCT 2053  
1981 ATCCAGGACACTGGGAGCACATAGAGATTACCCATGTTTGTGAACCTTAGAGTCATTCT 2040  
2054 CATGCTTTCTTTTATAATTCACATATATCCAGAGAAGATATGTTCTTGTAACTTGT 2113  
2041 CATGCTTTCTTTTATAATTCACATATATCCAGAGAAGATATGTTCTTGTAACTTGT 2100  
2114 ATACAACATAGCCCCAAATATAGTAAGATCTATCTAGATTAATCTAGATCAAAATGTTAG 2173  
2101 ATACAACATAGCCCCAAATATAGTAAGATCTATCTAGATTAATCTAGATCAAAATGTTAG 2160  
2174 AGATGCTATTTGATACAACTGTGGCCATGACTAGGAAAAGAGCTCAGCCCGCAGAGCTG 2233  
2161 AGATGCTATATGATACAACTGTGGCCATGACTAGGAAAAGAGCTCAGCCCGCAGAGCTG 2220  
2234 GGCTGCTCTCCGGAGGCGCAAAACCAAGAGGCTCTGCAAAAGTCAGGCTCAGGAGACTC 2293  
2221 GGCTGCTCTCCGGAGGCGCAAAACCAAGAGGCTCTGCAAAAGTCAGGCTCAGGAGACTC 2280  
2294 TGCCCTGCTCAGACCTCGGTGTGGACACACGCTGCATAGAGCTCTCTTGAACACAGAG 2353  
2281 TGCCCTGCTCAGACCTCGGTGTGGACACACGCTGCATAGAGCTCTCTTGAACACAGAG 2340  
2354 GGGTCTCAAGACATCTGCTTACCTATTAGCTTTCTTTTAACTTTTGGGG 2413  
2341 GGGTCTCAAGACATCTGCTTACCTATTAGCTTTCTTTTAACTTTTGGGG 2400  
2414 GAAAAGTATTTTGAAGAGTTTGTCTTGCATGATTTTATAAATAGTAAATPAAAGTTTTT 2473  
2401 GAAAAGTATTTTGAAGAGTTTGTCTTGCATGATTTTATAAATAGTAAATPAAAGTTTTT 2460  
2474 ACCATTAAAAAARAAAAA 2494  
2461 ACCATTAAAAAARAAAAA 2481

RESULT 2  
109499  
LOCUS 109499  
DEFINITION Sequence 20 from Patent WO 8909268.  
ACCESSION 109499  
VERSION 109499.1 GI:587805  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 2549)  
AUTHORS Baxter,R.C. and Wood,W.I.  
TITLE PRODUCTION OF INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN  
JOURNAL Patent: WO 8909268-A 20 05-OCT-1989;  
FEATURES Location/Qualifiers  
1. 2549  
source /organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 98.5%; Score 2457.4; DB 6; Length 2549;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2480; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

12 CAGCGCTTCTGCTGGATTCCACAGCTTCGCGCGTGTACTGTCCGCCCATCCCTGCG 71  
10 CAGCGCTTCTGCTGGATTCCACAGCTTCGCGCGTGTACTGTGCGCCCATCCCTGCG 69  
72 CGCCAGCCTGCCAAGCAGCGTGCCTGGTTGTCAGGGGTTCATGAGGGGGCGGACCCAC 131  
70 CGCCAGCCTGCCAAGCAGCGTGCCTGGTTGTCAGGGGTTCATGAGGGGGCGGACCCAC 129





2288 TCTGCCCTGCTGCAGACCTCGGTGTGGACACACGCTGCATAGAGCTCTCTCTGAAACAG 2347  
2352 AGGGGCTCAAGACATTCGCTACCTATTTAGCTTTCTTTTATTTTAACTTTTGG 2411  
2348 AGGGTCTCAAGACATTCGCTACCTATTTAGCTTTCTTTTATTTTAACTTTTGG 2407  
2412 GGGAAAAGTATTTTGAAGATTTGCTTGCATGTATTTATTAATAGTAAATAAAGTTT 2471  
2408 GGGAAAAGTATTTTGAAGATTTGCTTGCATGTATTTATTAATAGTAAATAAAGTTT 2467  
2472 TTACCATTAATAAAAAAAAAA 2494  
2468 TTACCATTAATAAAAAAAAAA 2490  
  
RESULT 3  
AX281786  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. 2496  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 253550.14"  
  
ORIGIN  
Query Match 98.5%; Score 2456; DB 6; Length 2496;  
Best Local Similarity 99.9%; Pred No. 0;  
Matches 2478; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
  
2y 1 GCGTGAGGATCAGCCGCTTCTGCTGATTCACAGCTTCGCGCGGTGATCTGCGC 60  
Db 18 GCGTGAGGATCAGCCGCTTCTGCTGATTCACAGCTTCGCGCGGTGATCTGCGC 77  
  
2y 61 CCATCCCTGCGCGCGCGCTGCCAA-GCAGCGGCCCGGTTGCAGGCGTCAAGCG 119  
Db 78 CCATCCCTGCGCGCGCGCGCTGCCAAAGCGCGCGCGGTTGCAGGCGTCAAGCG 137  
  
2y 120 GCGCGGACCAACGCTCTGGCGCTGCGCTGACTCTGCTGCTGCTGCTGCTGCTGCGCGCGCGC 179  
Db 138 GCGCGGACCAACGCTCTGGCGCTGCGCTGACTCTGCTGCTGCTGCTGCTGCTGCGCGCGCGC 197  
  
2y 180 GGTGGCGGCGCTGCGCGAGCTGCGGCGCTGGGTCGGCTGCGCTGCGCGCGCGCTG 239  
Db 198 GGTGGCGGCGCTGCGCGAGCTGCGGCGCTGGGTCGGCTGCGCTGCGCGCGCGCTG 257  
  
2y 240 CGAGCGCGCTGCACTGGCGGAGTGGCGCTTCGCGCGCGCTGCGCGCGCGCTGCGCGCGCTG 299  
Db 258 CGAGCGCGCTGCACTGGCGGAGTGGCGCTTCGCGCGCGCTGCGCGCGCGCTGCGCGCGCTG 317  
  
2y 300 CGAGCGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 359  
Db 318 CGAGCGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 377  
  
2y 360 CTACACCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 419  
Db 378 CTACACCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437  
  
2y 420 GCTGCGGCGCTGCTGAGCGCGCGCGCTCTGCGTCAAAGCTAGTCCGCTCAGCGCGCT 479

438 GCTGAGGCGCTGCTGAGCGCCGCGGCTCTGGTCAACGCTAGTCCGCTCAGCCGCT 497  
480 GCGGCGCTTACCTGCTGCGAGCGCCGCGCTCAGGAAATGCTAGTGTGCGAGGAGA 539  
498 GCGGCGCTTACCTGCTGCGAGCGCCGCGCTCAGGAAATGCTAGTGTGCGAGGAGA 557  
540 GCGAGCGCGCGAGTGTGGAGCGCGCTCCGCTCCAGCAGCGCCGCGGTGCTGATCC 599  
558 GCGAGCGCGCGAGTGTGGAGCGCGCTCCGCTCCAGCAGCGCCGCGGTGCTGATCC 617  
600 CAAAGTCCACCCCTCCATTCAGAGATAATCATCATCAAGAAAGGCGATGCTAAAGACAG 659  
618 CAAAGTCCACCCCTCCATTCAGAGATAATCATCATCAAGAAAGGCGATGCTAAAGACAG 677  
660 CAGGCGCTACAAAGTGTGAGTCTCAGAGCACAGATACCCAGAACTTCTCTCCGA 719  
678 CCAGCGCTACAAAGTGTGAGTCTCAGAGCACAGATACCCAGAACTTCTCTCCGA 737  
720 GTCCAAAGCGGAGACAGAAATATGCTCCGCTGAGAGAAATGGAAGACACACTGAATCA 779  
738 GTCCAAAGCGGAGACAGAAATATGCTCCGCTGAGAGAAATGGAAGACACACTGAATCA 797  
780 CCGAAGTTCCTCAATGTGCTGAGTCCAGGCGGTGACATTCCTCAACTGTGACAGAA 839  
798 CCGAAGTTCCTCAATGTGCTGAGTCCAGGCGGTGACATTCCTCAACTGTGACAGAA 857  
840 GGGATTTTATAGAAAGACAGTGTGCGCCCTTCCAAAGGCGAGGAGCGGCTTCTGCTG 899  
858 GGGATTTTATAGAAAGACAGTGTGCGCCCTTCCAAAGGCGAGGAGCGGCTTCTGCTG 917  
900 GTGTGCTGATTAAGTATGGGCGAGCTCTCCAGGCTACACCAAGGCGGAGGAGGAGCT 959  
918 GTGTGCTGATTAAGTATGGGCGAGCTCTCCAGGCTACACCAAGGCGGAGGAGGAGCT 977  
960 GCACCTGTACAGCATGACAGCAAGTACAGCTCCGCGCAAGGTTAAATGTGGAGCTCAAA 1019  
978 GCACCTGTACAGCATGACAGCAAGTACAGCTCCGCGCAAGGTTAAATGTGGAGCTCAAA 1037  
1020 TATGCTTTATTTGCAAAAAGATGCGCAAGGACATGACAGCGAGCTGGCTACAGCTCG 1079  
1038 TATGCTTTATTTGCAAAAAGATGCGCAAGGACATGACAGCGAGCTGGCTACAGCTCG 1097  
1080 ATTATATTTCTGTTGTTGGTGAATTTTTTTTAAACCAAGTTTAAAGAGAGTT 1139  
1098 ATTATATTTCTGTTGTTGGTGAATTTTATTTTAAACCAAGTTTAAAGAGAGTT 1156  
1140 TTTGAAATGCTTATGTTTCTTTGAATGGTAAACTTGAGCATCTTTTCACTTCCAGTAG 1199  
1157 TTTGAAATGCTTATGTTTCTTTGAATGGTAAACTTGAGCATCTTTTCACTTCCAGTAG 1216  
1200 TCAGCAAGAGCGAGTTGAATTTTCTGCTTCTTATCAAAATATTCAGAGACTCGAG 1259  
1217 TCAGCAAGAGCGAGTTGAATTTTCTGCTTCTTATCAAAATATTCAGAGACTCGAG 1276  
1260 CACAGCACCCAGACTTCATGCGCGCGTGAATGTCCACCATGTTGGTCAAGCGCGC 1319  
1277 CACAGCACCCAGACTTCATGCGCGCGTGAATGTCCACCATGTTGGTCAAGCGCGC 1336  
1320 ACCACTGATTTGATCTAGCGGCTGTGTTGCTATGTAGAGACACGCTTCAACCCC 1379  
1337 ACCACTGATTTGATCTAGCGGCTGTGTTGCTATGTAGAGACACGCTTCAACCCC 1396  
1380 ACTCCCGGTACAGTGCAGCAGGCTTTATCGAGATAGGAAACCTTTAAACCCCGGTCA 1439  
1397 ACTCCCGGTACAGTGCAGCAGGCTTTATCGAGATAGGAAACCTTTAAACCCCGGTCA 1456  
1440 TCGGAGCATCCAAAGCATGCTCTGAGCTCAAGCCTTCTGTGGTGTCTATTTCTGAAA 1499  
1457 TCGGAGCATCCAAAGCATGCTCTGAGCTCAAGCCTTCTGTGGTGTCTATTTCTGAAA 1516  
1500 CAAAGGCGTGGATCCCTCAACCAAGAGAAATGTTTATGTTTCAAGTACCTGTACTGCT 1559

1517	CAAGGCGTGGATCCTCAACCAAGGAAGATGTTATATGCTTCAAGTGACCTGTACTGCT	1576
1560	TGGGGACTATTGGAGAAAATAAGGTGGAGCTCTACTTGTGTTAAAAAATATGTATCTAAGA	1619
1577	TGGGGACTATTGGAGAAAATAAGGTGGAGCTCTACTTGTGTTAAAAAATATGTATCTAAGA	1636
1620	ATGTTCTAGGGCACTCTGGGAACCTATAAGGCAGGTATTTGGGGCCCTCCCTCTTCAGGA	1679
1637	ATGTTCTAGGGCACTCTGGGAACCTATAAGGCAGGTATTTGGGGCCCTCCCTCTTCAGGA	1696
1680	ATCTTCTCGAAGACATCGGCCAGTCGAAGGCCCCAGGATGGCTTTTGTCTCGGGCCCCGTGG	1739
1697	ATCTTCTCGAAGACATCGGCCAGTCGAAGGCCCCAGGATGGCTTTTGTCTCGGGCCCCGTGG	1756
1740	GGTAGAGGGACAGAGACACAGGGAGAGTCAGCCTCCACATTCAGAGGCATCACAGTAA	1799
1757	GGTAGAGGGACAGAGACACAGGGAGAGTCAGCCTCCACATTCAGAGGCATCACAGTAA	1816
1800	TGSCACAAATCTTTCGGATGATCTGCAGAAAATAGTGTGTTTTGTAGTTTCAACAACCTCAAGACG	1859
1817	TGSCACAAATCTTTCGGATGATCTGCAGAAAATAGTGTGTTTTGTAGTTTCAACAACCTCAAGACG	1876
1860	AAGCTTATTCTGAGGATAAGCTCTTTAAAGGCAAGCTTTATTTTCACTCTCTCACTTT	1919
1877	AAGCTTATTCTGAGGATAAGCTCTTTAAAGGCAAGCTTTATTTTCACTCTCTCACTTT	1936
1920	TGTCCTCTTTAGCACAAATGTAAAAAGAAATAGTAATATCAGAACAGGAAGGAGGAATGGC	1979
1937	TGTCCTCTTTAGCACAAATGTAAAAAGAAATAGTAATATCAGAACAGGAAGGAGGAATGGC	1996
1980	TTGCTGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTTGTGAA	2039
1997	TTGCTGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTTGTGAA	2056
2040	CTT*AGAGTCATTCTCATGCTTTCTTTATATAATCACAATATATGCAGAGAAGATATGTT	2099
2057	CTTAGAGTCATTCTCATGCTTTCTTTATATAATCACAATATATGCAGAGAAGATATGTT	2116
2100	CTTGTTAACTATTGTAATACACATAGCCCCAAAATATAGTAAGATCTATCTAGATAATCCT	2159
2117	CTTGTTAACTATTGTAATACACATAGCCCCAAAATATAGTAAGATCTATCTAGATAATCCT	2176
2160	AGATGAATGTTAGAGATGCTATTTGATCAAACTGTGCGCATGACATGAGGAAGGAGCTC	2219
2177	AGATGAATGTTAGAGATGCTATTTGATCAAACTGTGCGCATGACATGAGGAAGGAGCTC	2236
2220	ACGCCACAGACATGGGTGCTCTCTCCGGAGGCCAAACCCAAAGAGGTCCTGGCAAAAGTCAG	2279
2237	ACGCCACAGACATGGGTGCTCTCTCCGGAGGCCAAACCCAAAGAGGTCCTGGCAAAAGTCAG	2296
2280	GCTCAGGGAGACTCTGCCCTGCTCGACACTCGGTGTGGACACACGCTGCATAGAGCTCT	2339
2297	GCTCAGGGAGACTCTGCCCTGCTCGACACTCGGTGTGGACACACGCTGCATAGAGCTCT	2356
2340	CCTTGAAAAACAGAGGGTCTCAAGACATTTCTGCCCTAGCTTTTCTTTTATTTTTT	2399
2357	CCTTGAAAAACAGAGGGTCTCAAGACATTTCTGCCCTAGCTTTTCTTTTATTTTTT	2416
2400	TAACCTTTTGGGGGAAAGTATTTTTGAGAAGTTGTCCTTGCAATGTATTTATAAATAG	2459
2417	TAACCTTTTGGGGGAAAGTATTTTTGAGAAGTTGTCCTTGCAATGTATTTATAAATAG	2476
2460	TAAATAAAGTTTTTACCATT	2479
2477	TAAATAAAGTTTTTACCATT	2496

insulin-like growth factor-binding protein 3.  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 2499)  
 Thweatt,R., Fleischmann,R. and Goldstein,S.  
 Analysis of the primary structure of insulin-like growth factor  
 binding protein-3 cDNA from Werner syndrome fibroblasts  
 DNA Seq. 4 (1), 43-46. (1993)  
 94146408  
 7508771  
 Thweatt,R.  
 Direct Submission  
 Submitted (09-MAR-1992) R. Thweatt, Univ Arkansas for Med Sciences,  
 John L McClellan VA, 4300 West 7th Street/Res 151, Little Rock, AR  
 72205, USA

FEATURES  
 Location/Qualifiers  
 1..2499  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /isolates="Werners syndrome subject TO"  
 /db\_xref="taxon:9606"  
 /cell\_line="WS8"  
 /cell\_type="fibroblast"  
 /issue\_type="skin"  
 /clone\_lib="Werners syndrome fibroblast cDNA"  
 /dev\_stage="adult"  
 1..2499  
 /gene="IGFBP-3"  
 133..1009  
 /gene="IGFBP-3"  
 /codon\_start=1  
 /product="insulin-like growth factor binding protein 3"  
 /protein\_id="CA846087.1"  
 /db\_xref="GI:396164"  
 /db\_xref="GOA:PI7936"  
 /db\_xref="SWISS-PROT:PI7936"  
 /translation="MCQARPTLWAAALTLLVLRGPPVARAGASSGGSLGVFVRCPECD  
 ARALAQAPPAPVACAVLREPGGCGTLCALSEGQPCGIYTERCGSLRCQPSDEAR  
 ALIQAQDGLCLVNAISAVLRALYLLPAPAGNASSEDRSAGSVESFSVSTRVA  
 SDFKPEHLPSKIIILKKHAKDSQRVKVDYEOSITQNFSESSEKRETYGPPRRME  
 DTNLHLKLVNLSPRGVHPNCDKGFYKKQCRPSKGRKRGFCWCVDKYGQPLPGYV  
 TKGEDVHCYSMQSK"  
 133..213  
 /gene="IGFBP-3"  
 2482..2486  
 /gene="IGFBP-3"

sig\_peptide  
 polyA\_signal

ORIGIN  
 Query Match 98.1%; Score 2446.2; DB 9; Length 2499;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 2465; Conservative 0; Mismatches 13; Indels 1; Gaps 1;  
 Qy 1 GCGGTGAGGATCAGCCGCTTCCTGCTGGATTCCACAGCTTGGCGCGGTGACTGTGCC 60  
 Db 22 GCGGTGAGGATCAGCCGCTTCCTGCTGGATTCCACAGCTTGGCGCGGTGACTGTGCC 81  
 Qy 61 CCATCCCTGGCGCCAGCTGCCAAGCAGCTGCCCGGTTGCAGCGCTCATGCAGCGG 120  
 Db 82 CCATCCCTGGCGCCAGCTGCCAAGCAGCTGCCCGGTTGCAGCGCTCATGCAGCGG 141  
 Qy 121 GCGGACCCACGCTCTGGGCGGTGCCTGACTCTGTGTGCTCTCCGGGGCGCGCG 180  
 Db 142 GCGGACCCACGCTCTGGGCGGTGCCTGACTCTGTGTGCTCTCCGGGGCGCGCG 201  
 Qy 181 GTGCGGGGCTGCGGAGCTCGGGGGGCTTGGGTCCCGTGGTGGCTGTGCGAGCCGTCG 240  
 Db 202 GTGCGGGGCTGCGGAGCTCGGGGGGCTTGGGTCCCGTGGTGGCTGTGCGAGCCGTCG 261  
 Qy 241 GACGCGGTGCACTGGCCCAAGTGCGCGCCCTCCGCGCGCGGTGTGCGGAGCTGGTGGCG 300



QY 2461 AAATAAGTTTTTACCATT 2479  
Db 2481 AAATAAGTTTTTACCATT 2499

RESULT 5  
LOCUS AR364291 2550 bp DNA linear PAT 03-SEP-2003  
DEFINITION Sequence 23 from patent US 5258287.  
ACCESSION AR364291  
VERSION AR364291.1 GI:34426677  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2550)  
AUTHORS Baxter,R.C. and Wood,W.I.  
TITLE DNA encoding and methods of production of insulin-like growth factor binding protein Bp3  
JOURNAL Patent: US 5258287-A 23 02-NOV-1993;  
FEATURES  
Location/Qualifiers  
1..2550  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 98.1%; Score 2446.2; DB 6; Length 2550;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 2473; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 12 CAGCGGCTTCCTGCTGATTCACAGCTTCGCGCGCTGACTGTCGCCCATCCCTGCG 71  
Db 10 CAGCGGCTTCCTGCTGATTCACAGCTTCGCGCGCTGACTGTCGCCCATCCCTGCG 69

QY 72 CGCCAGGCTCCAGAGAGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 131  
Db 70 CGCCAGGCTCCAGAGAGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 129

QY 132 GCTCTGGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 191  
Db 130 GCTCTGGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189

QY 192 TGGCGGAGCTCGGGGGCTTGGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 251  
Db 190 TGGCGGAGCTCGGGGGCTTGGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 249

QY 252 ACTGGCCAGTGGCGCTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 311  
Db 250 ACTGGCCAGTGGCGCTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309

QY 312 CGGCTGCTGCTGAGTGGCACTGAGCGAGGCGAGCGCTGCGGCACTACACCGAGCG 371  
Db 310 CGGCTGCTGCTGAGTGGCACTGAGCGAGGCGAGCGCTGCGGCACTACACCGAGCG 369

QY 372 CTGTGGCTCGGCGCTTCGCTGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTC 431  
Db 370 CTGTGGCTCGGCGCTTCGCTGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTC 429

QY 432 GCTGAGCGGCGCGGCTTCGCTGCAACGCTAGTCCGCTGAGCGCGCTGCGCGCTTACCT 491  
Db 430 GCTGAGCGGCGCGGCTTCGCTGCAACGCTAGTCCGCTGAGCGCGCTGCGCGCTTACCT 489

QY 492 GCTGCGAGGCGCGGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCG 551  
Db 490 GCTGCGAGGCGCGGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCG 549

QY 552 CAGTGTGAGAGCGCGCTTCGCTGCAACGCTAGTCCGCTGAGCGCGCTGCGCGCTTACC 611  
Db 550 CAGTGTGAGAGCGCGCTTCGCTGCAACGCTAGTCCGCTGAGCGCGCTGCGCGCTTACC 609

QY 612 CTTCCATTCAAGATAATCATCATCAAGAGGCGATGCTAAAGACAGCGCGCTACAA 671  
Db 610 CTTCAATTCAAGATAATCATCATCAAGAGGCGATGCTAAAGACAGCGCGCTACAA 669

672 AGTTGACTACGAGTCTCAGAGCACAGATACCAGAACTTCTCTCCGAGTCCAAAGCGGA 731  
Db 670 AGTTGACTACGAGTCTCAGAGCACAGATACCAGAACTTCTCTCCGAGTCCAAAGCGGA 729

QY 732 GACAGAATATGGTCCCTGCTAGAGAAATCGAGACACACTGATACCTCGAAGTTCCT 791  
Db 730 GACAGAATATGGTCCCTGCTAGAGAAATCGAGACACACTGATACCTCGAAGTTCCT 789

QY 792 CAATGTGCTGAGTCCAGGGGTGTACACATTCCTCAACTGTGACAAAGAGGATTTTATA 851  
Db 790 CAATGTGCTGAGTCCAGGGGTGTACACATTCCTCAACTGTGACAAAGAGGATTTTATA 849

QY 852 GAAAAAGCAGTGTGCGCTTCCTCAAGGCGAGAGCGGGCTTCTGCTGTGTGGATAA 911  
Db 850 GAAAAAGCAGTGTGCGCTTCCTCAAGGCGAGAGCGGGCTTCTGCTGTGTGGATAA 909

QY 912 GTATGGGCGAGCTCTCCAGGCTACACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 971  
Db 910 GTATGGGCGAGCTCTCCAGGCTACACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 969

QY 972 CATGAGAGCAAGTAGAGCGCTGCGCGAAGGTATATGTGGAGCTCAATATATGCTTTT 1031  
Db 970 CATGAGAGCAAGTAGAGCGCTGCGCGAAGGTATATGTGGAGCTCAATATATGCTTTT 1029

QY 1032 TGCACAAAAGACTGCGCAAGGACATGACAGCAGCTGGCTACAGCCTCGATTTATATTTCT 1091  
Db 1030 TGCACAAAAGACTGCGCAAGGACATGACAGCAGCTGGCTACAGCCTCGATTTATATTTCT 1089

QY 1092 GTTGTGTGAACTGATTTTTTTTAAACCAAGTTTAAAGAGGTTTGTGAAATGCT 1151  
Db 1090 GTTGTGTGAACTGATTTTTTTTAAACCAAGTTTAAAGAGGTTTGTGAAATGCT 1148

QY 1152 ATGTTTCTTTTGAATGTTAACTTGCAGCATCTTTTCACTTTTCCAGTAGTCAGCAAGAGC 1211  
Db 1149 ATGTTTCTTTTGAATGTTAACTTGCAGCATCTTTTCACTTTTCCAGTAGTCAGCAAGAGC 1208

QY 1212 AGTTGAAATTTTGTGCTGCTTATCAAAATATTCAGAGACTCGAGCAGACGCCAG 1271  
Db 1209 AGTTGAAATTTTGTGCTGCTTATCAAAATATTCAGAGACTCGAGCAGACGCCAG 1268

QY 1272 ACTTCATGCGCGCGGGAATGCTCACCACATGTTGGTGAAGCGGCCACCACTGACTTT 1331  
Db 1269 ACTTCATGCGCGCGGGAATGCTCACCACATGTTGGTGAAGCGGCCACCACTGACTTT 1328

QY 1332 GTGACTTAGCGGCTGTGTGCTATGTAGAGAACACGCTTACCCCCACTCCCCGCTACA 1391  
Db 1329 GTGACTTAGCGGCTGTGTGCTATGTAGAGAACACGCTTACCCCCACTCCCCGCTACA 1388

QY 1392 GTGCGCAGAGGCTTTATCGAGNATAGGAAACCTTTAACCCTCGTATCCGGACATCCC 1451  
Db 1389 GTGCGCAGAGGCTTTATCGAGNATAGGAAACCTTTAACCCTCGTATCCGGACATCCC 1448

QY 1452 AACGATGCTCCTGGAGCTCAGAGCTTCTGTGTGTCATTTCTGAAACAAAGGGCTGGA 1511  
Db 1449 AACGATGCTCCTGGAGCTCAGAGCTTCTGTGTGTCATTTCTGAAACAAAGGGCTGGA 1508

QY 1512 TCCCTCAACCAAGAGATGTTATGCTCTCAAGTGAAGTGTGCTGCTGCTGCTGCTGCTG 1571  
Db 1509 TCCCTCAACCAAGAGATGTTATGCTCTCAAGTGAAGTGTGCTGCTGCTGCTGCTGCTG 1568

QY 1572 GAGAAAAATAGGTGAGTCTTACTTGTTTAAAAATATGTATCTAAGAAATGTTCTAGGCG 1631  
Db 1569 GAGAAAAATAGGTGAGTCTTACTTGTTTAAAAATATGTATCTAAGAAATGTTCTAGGCG 1628

QY 1632 ACTTGGGAACCTATAAGGAGGATTTTCGGGCCCTCTCTTTCAGGAATCTTCTCTGAG 1691  
Db 1629 ACTTGGGAACCTATAAGGAGGATTTTCGGGCCCTCTCTTTCAGGAATCTTCTCTGAG 1688

QY 1692 ACATGGCCCGAGTCGAAGGCCAGGATGCTTTTGTGCGGCCCGCTGGGGTAGGAGGAC 1751  
Db 1689 ACATGGCCCGAGTCGAAGGCCAGGATGCTTTTGTGCGGCCCGCTGGGGTAGGAGGAC 1748

```

1752 AGAGAGACAGGAGAGATCAGCTCCACATTCAGAGGATCATCAAGTAATGGCACAAATCT 1811
1749 AGAGAGAC-GGGAGAGTCAAGCTCCACATTCAGAGGATCATCAAGTAATGGCACAAATCT 1807
1812 TCGGATGATCGAGAAATAGTGTGTTGTTAGTTCACAACTCAAGCAGAGCTTATTTCT 1871
1808 TCGGATGATCGAGAAATAGTGTGTTGTTAGTTCACAACTCAAGCAGAGCTTATTTCT 1867
1872 GAGGATAAGCTCTTTAAAGGCAAAAGCTTTATTTTCACTCTCATCTTTGCTCTCTTAG 1931
1868 GAGGATAAGCTCTTTAAAGGCAAAAGCTTTATTTTCACTCTCATCTTTGCTCTCTTAG 1927
1932 CACAATGTAAGAAAGATAGTATATATCAGAACAGGAGAGATGCTGCTGGGGAGC 1991
1928 CACAATGTAAGAAAGATAGTATATATCAGAACAGGAGAGATGCTGCTGGGGAGC 1987
1992 CCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTTGGTGAAGCTAGAGTCATT 2051
1988 CCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTTGGTGAAGCTAGAGTCATT 2047
2052 CTGATGCTTTCTTTTATATATCACAATATATGAGAGAGATATGTTCTGTTTAACTT 2111
2048 CTGATGCTTTCTTTTATATATCACAATATATGAGAGAGATATGTTCTGTTTAACTT 2107
2112 GTATACACATAGCCCAATATAGTAAGATCTATCTAGATAATCCTAGATAAATGTT 2171
2108 GTATACACATAGCCCAATATAGTAAGATCTATCTAGATAATCCTAGATAAATGTT 2167
2172 AGAGATGCTATTTGATACAACTGTGCGCATGCTAGGAAAGAGCTACGCCAGAGAC 2231
2168 AGAGATGCTATATGATACAACTGTGCGCATGCTAGGAAAGAGCTACGCCAGAGAC 2227
2232 TGGGCTGCTCTCCGAGGCGCAACCCAGAGAGCTGCGCAAGTCAAGCTCAGGAGAC 2291
2228 TGGGCTGCTCTCCGAGGCGCAACCCAGAGAGCTGCGCAAGTCAAGCTCAGGAGAC 2287
2292 TCTGCGCTGTGTCAGACCTCGGTGTGCGACACACGCTGCTAGAGCTCTCCTTGAAACAG 2351
2288 TCTGCGCTGTGTCAGACCTCGGTGTGCGACACACGCTGCTAGAGCTCTCCTTGAAACAG 2347
2352 AGGGTCTGAGACATCTGCTACCTATAGCTTCTTTTATTTTAACTTTTGGG 2411
2348 AGGGTCTGAGACATCTGCTACCTATAGCTTCTTTTATTTTAACTTTTGGG 2407
2412 GGGAAAGTATTTTGAAGATTTGTTGCAATGATTTTAAATAGTAATAAAGTTT 2471
2408 GGGAAAGTATTTTGAAGATTTGTTGCAATGATTTTAAATAGTAATAAAGTTT 2467
2472 TTACCAATATAAAAAAAAAAAAAA 2494
2468 TTACCAATATAAAAAAAAAAAAAA 2490

RESULT 6
LOCUS BC018962
DEFINITION Homo sapiens insulin-like growth factor binding protein 3, mRNA
ACCESSION BC018962
VERSION BC018962.2
KEYWORDS G1:33879043
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2497)
REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Moradinejad, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

```

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 McKernan, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W.,  
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettelman, M., Madan, A.C., Rodrigues, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smallus, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2338257  
 12477932  
 2 (bases 1 to 2497)  
 Direct Submission  
 Submitted (07-DEC-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 On Aug 19, 2003 this sequence version replaced gi:17511987.  
 Contact: MGC help desk  
 Email: [cgaps-x@mail.nih.gov](mailto:cgaps-x@mail.nih.gov)  
 Tissue Procurement: DCTD/DTF  
 CNA Library Preparation: Rubin Laboratory  
 CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nigr.nih.gov](mailto:nisc_mgc@nigr.nih.gov)  
 Akhter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hagnigni, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W.,  
 Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 27 Row: c Column: 22.

FEATURES  
 Location/Qualifiers

```

1..2497
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:19695 IMAGE:3506929"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH MGC 14"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1..2497
/gene="IGFBP3"
/notes="synonym: IBP3"
/db_xref="LocusID:3486"
/db_xref="MIM:146732"
79..954
/codon_start=1
/product="insulin-like growth factor binding protein 3"
/protein_id="AAH18962.1"
/db_xref="GI:17511988"
/translaton="MORAPTLMAAATLLVLRGPVVARAGASSGGLGPVVRCEPCD
ARALAQCAPPPACVACELVRPCCGCTCALSCQPGCIYTERCGSLRQCPSPDEAR
PQALLDGRGICVNAVASVRLAYLLPAPAGNASESDRSAGSVESVSTHVR
SDPKFHLHSKILTIKKGAKDSORYKDYVESQSTDTQNFSSKRETEYGPCEHRE
DTLNLHLKFLNVLSPRGVHPCNDKXGFKKKQCRPSKRGKGFQWCVCVDKYQPLPGYT

```

gene

CDS

TKXEDVHCYSNQS<sup>190..426</sup>  
misc\_feature /note="IB; Region: Insulin growth factor-binding protein homologues"  
/db\_xref="CDD:smart00121"  
misc\_feature 715..933  
/note="Thyroglobulin\_1; Region: Thyroglobulin type-1 repeat. Thyroglobulin type 1 repeats are thought to be involved in the control of proteolytic degradation. The domain usually contains six conserved cysteines. These form three disulphide bridges. Cysteines 1 pairs with 2, 3 with 4 and 5 with 6"  
/db\_xref="CDD:pfam00086"

ORIGIN

Query Match 98.1%; Score 2445.8; DB 9; Length 2497;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2458; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 34 CACAGCTTCGGCGCGTGTCTGCTCCCATCCCTGCGCGCCAGCCTGCCAAGCAGCGT 93  
DB 1 CACAGCTTCGGCGCGTGTCTGCTCCCATCCCTGCGCGCCAGCCTGCCAAGCAGCGT 60  
QY 94 GCCCGGTTCAGGCGTCATGACAGGGGCGGACCCACGCTCTGGGCGCTGCGCTGACT 153  
DB 61 GCCCGGTTCAGGCGTCATGACAGGGGCGGACCCACGCTCTGGGCGCTGCGCTGACT 120  
QY 154 CTGCTGCTGCTCTCCGGGGCGCGCGGTGCGGGGCTGCGGAGCTGGGGGGCTTG 213  
DB 121 CTGCTGCTGCTCTCCGGGGCGCGCGGTGCGGGGCTGCGGAGCTGGGGGGCTTG 180  
QY 214 GGTCCCGTGTGCTGCGTGCAGCGTGTGCGACGCGGTGCATGCGCCAGTGGCGCCCTCCG 273  
DB 181 GGTCCCGTGTGCTGCGTGCAGCGTGTGCGACGCGGTGCATGCGCCAGTGGCGCCCTCCG 240  
QY 274 CCGCGGTGTGCGGAGCTGTGCGGAGCGGGCTGCGGCTGCGGCTGCTGAGCGGCGCA 333  
DB 241 CCGCGGTGTGCGGAGCTGTGCGGAGCGGGCTGCGGCTGCGGCTGCTGAGCGGCGCA 300  
QY 334 CTGAGCGAGGGCGACCGTGTGCGGCACTACACCGAGCGCTGTGCTCGGGCTTCGCTGC 393  
DB 301 CTGAGCGAGGGCGACCGTGTGCGGCACTACACCGAGCGCTGTGCTCGGGCTTCGCTGC 360  
QY 394 CAGCGTGTGCGGAGCGGACCGTGTGCGGCGCTGAGCGGCTGTGAGCGCGGGCTCTGC 453  
DB 361 CAGCGTGTGCGGAGCGGACCGTGTGCGGCGCTGAGCGGCTGTGAGCGCGGGCTCTGC 420  
QY 454 GTCAACGCTAGTGTGCGCTCAGCGCGCTGCGCGCTACCTGCTGCCAGCGCGCCAGCTCCA 513  
DB 421 GTCAACGCTAGTGTGCGCTCAGCGCGCTGCGCGCTACCTGCTGCCAGCGCGCCAGCTCCA 480  
QY 514 GGAATGTAGTGTGAGTGTGAGGAGAACCGGACCGCGCGCTGAGAGCGCGCTCGTC 573  
DB 481 GGAATGTAGTGTGAGTGTGAGGAGAACCGGACCGCGCGCTGAGAGCGCGCTCGTC 540  
QY 574 TCCAGCAGCCAGCGGTGTCTGATCCCAAGTTCACCCCTCCATTCACCAAGATATCATC 633  
DB 541 TCCAGCAGCCAGCGGTGTCTGATCCCAAGTTCACCCCTCCATTCACCAAGATATCATC 600  
QY 634 ATCAAGAAAGGCGCTGTAAAGACGCGGCTCAAAAGTTGACTACGAGTCTCAGAGC 693  
DB 601 ATCAAGAAAGGCGCTGTAAAGACGCGGCTCAAAAGTTGACTACGAGTCTCAGAGC 660  
QY 694 ACAGATACCCAGAACTTCTCTCCAGTCCCAAGCGGGAGACAGAAATGTGTCCTGCGGT 753  
DB 661 ACAGATACCCAGAACTTCTCTCCAGTCCCAAGCGGGAGACAGAAATGTGTCCTGCGGT 720  
QY 754 AGAGAAATGGAGACACACTGAAATCACTGAAAGTTCCCTCAATGTGTGAGTCCAGGGGT 813  
DB 721 AGAGAAATGGAGACACACTGAAATCACTGAAAGTTCCCTCAATGTGTGAGTCCAGGGGT 780  
QY 814 GTACACATTCCTCAACTGTGACAAAGAGGATTTTATAAGAAAAAGCAGTGTGCGCCCTCC 873

DB 781 GTACACATTCCTCCAACTGTGACAAAGAGGATTTTATAAGAAAAAGCAGTGTGCGCCCTCC 840  
QY 874 AAAGCAGGAAGCGGGCTTCGCTGGTGTGGATAAGTATGGCAGCCTCTCCAGGC 933  
DB 841 AAAGCAGGAAGCGGGCTTCGCTGGTGTGGATAAGTATGGCAGCCTCTCCAGGC 900  
QY 934 TACACCAACAAGGGAAGGAGGACGTCGACCTGTACACATGCAGCATGCAGCAAGTAGACGCT 993  
DB 901 TACACCAACAAGGGAAGGAGGACGTCGACCTGTACACATGCAGCATGCAGCAAGTAGACGCT 960  
QY 994 GCCGCAAGGTTAATGTGGAGCTCAAAATATGCCTTATTTTGCAAAAAAGACTGCCAAGGAC 1053  
DB 961 GCCGCAAGGTTAATGTGGAGCTCAAAATATGCCTTATTTTGCAAAAAAGACTGCCAAGGAC 1020  
QY 1054 ATGACCAAGCAGCTGGCTACAGCCTCGATTTATTTCTGTTGTGCTGAACGATTTTTT 1113  
DB 1021 ATGACCAAGCAGCTGGCTACAGCCTCGATTTATTTCTGTTGTGCTGAACGATTTT 1079  
QY 1114 TTTAAACCAAGGTTTAGAAGAGGTTTTTGAATGCCTATGCTTTCTTTGAATGTAAC 1173  
DB 1080 TTTAAACCAAGGTTTAGAAGAGGTTTTTGAATGCCTATGCTTTCTTTGAATGTAAC 1139  
QY 1174 TTGAGCATCTTTCACTTTCCAGTAGTCAGCAAGAGCAGTTTGAATTTCTTGTGCTT 1233  
DB 1140 TTGAGCATCTTTCACTTTCCAGTAGTCAGCAAGAGCAGTTTGAATTTCTTGTGCTT 1199  
QY 1234 CCTATCAAAATATTCAGAGACTCGAGCAGCAGCCAGACTTTCATGCGCCCTGGAATGC 1293  
DB 1200 CCTATCAAAATATTCAGAGACTCGAGCAGCAGCCAGACTTTCATGCGCCCTGGAATGC 1259  
QY 1294 TCACACATGTTGGTGCAGAGCGGCGCACCTGACTTTGTGACTTTAGGGGGCTGTGTC 1353  
DB 1260 TCACACATGTTGGTGCAGAGCGGCGCACCTGACTTTGTGACTTTAGGGGGCTGTGTC 1319  
QY 1354 CTATGTAGAGAACACGCTTCACCCCATCTCCCGTACAGTGCGCACAGGCTTTATCGAGA 1413  
DB 1320 CTATGTAGAGAACACGCTTCACCCCATCTCCCGTACAGTGCGCACAGGCTTTATCGAGA 1379  
QY 1414 ATAGGAAAACTTTAAACCCCGGTATCGGACATCCCAACGCAATGCTCTGAGAGCTCAC 1473  
DB 1380 ATAGGAAAACTTTAAACCCCGGTATCGGACATCCCAACGCAATGCTCTGAGAGCTCAC 1439  
QY 1474 AGCTTCTGTGTGTCTATTTCTGAAACAAAGGGCGTGGATCCCTCAACCAAGAAAGATGT 1533  
DB 1440 AGCTTCTGTGTGTCTATTTCTGAAACAAAGGGCGTGGATCCCTCAACCAAGAAAGATGT 1499  
QY 1534 TATGCTTCAAGTACCTGTACTGCTTGGGACTATTTGGAGAAAAATAAGGTGAGTCTCTA 1593  
DB 1500 TATGCTTCAAGTACCTGTACTGCTTGGGACTATTTGGAGAAAAATAAGGTGAGTCTCTA 1559  
QY 1594 CTGTTTAAABAATATGTATCTAAGAAATGTTCTAGGGCACTCTGGGAACCTATAAGGCA 1653  
DB 1560 CTGTTTAAABAATATGTATCTAAGAAATGTTCTAGGGCACTCTGGGAACCTATAAGGCA 1619  
QY 1654 GGTATTTGCGGGCTCTCTCTTCAGGAATCTTCTCGAAGACATGCGCCAGTGGAGGCCA 1713  
DB 1620 GGTATTTGCGGGCTCTCTCTTCAGGAATCTTCTCGAAGACATGCGCCAGTGGAGGCCA 1679  
QY 1714 GATGCTTTCGTCGGGCGCGTGGGTAGAGGAGCAGAGAGCAGGAGAGTCAGC 1773  
DB 1680 GATGCTTTCGTCGGGCGCGTGGGTAGAGGAGCAGAGAGCAGGAGAGTCAGC 1739  
QY 1774 TCCACATTCAGAGGATCAACAAGTAATGSCACAAATTTCTTGGATGACTGCAGAAAAATAGT 1833  
DB 1740 TCCACATTCAGAGGATCAACAAGTAATGSCACAAATTTCTTGGATGACTGCAGAAAAATAGT 1799  
QY 1834 GTTTTGTAGTTCACAACTCAAGAGCGAGCTTATTTCTGAGGATAAGCTCTTTAAAGCA 1893  
DB 1800 GTTTTGTAGTTCACAACTCAAGAGCGAGCTTATTTCTGAGGATAAGCTCTTTAAAGCA 1859  
QY 1894 AGCTTTATTTTCTCATCTCTCATCTTTTGTCTCTCTCTAGCACAATGTAAAAAGAAATAGTA 1953  
DB 1860 AGCTTTATTTTCTCATCTCTCATCTTTTGTCTCTCTCTAGCACAATGTAAAAAGAAATAGTA 1919



```

1954 ATATCAGAACAGAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGACACTGGGAGCAC 2013
1920 ATATCAGAACAGAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGACACTGGGAGCAC 1979
2014 ATAGAGATTACCCATGTTTGTGAATCTAGAGTCACTCTGATGCTTCTTTTATAATTC 2073
1980 ATAGAGATTACCCATGTTTGTGAATCTAGAGTCACTCTGATGCTTCTTTTATAATTC 2039
2074 ACACATATATGACAGAGAGATATGTTTCTTGTAAATCTAGATGATGCTTCTTTTATAATTC 2133
2040 ACACATATATGACAGAGAGATATGTTTCTTGTAAATCTAGATGATGCTTCTTTTATAATTC 2099
2134 TAGTAAGATCTATATCTAGATGATGCTTCTTGTAAATCTAGATGATGCTTCTTTTATAATTC 2193
2100 TAGTAAGATCTATATCTAGATGATGCTTCTTGTAAATCTAGATGATGCTTCTTTTATAATTC 2159
2194 GTGGCCATGACTGAGAGAGAGAGCTCAGCCAGAGAGAGCTGCTCTCCGAGAGGCA 2253
2160 GTGGCCATGACTGAGAGAGAGAGCTCAGCCAGAGAGAGCTGCTCTCCGAGAGGCA 2219
2254 AACCCAGAGAGCTGAGAGAGAGAGCTCAGCCAGAGAGAGCTGCTCTCCGAGAGGCA 2313
2220 AACCCAGAGAGCTGAGAGAGAGAGCTCAGCCAGAGAGAGCTGCTCTCCGAGAGGCA 2279
2314 TGTGGACACAGCTGATAGAGCTCTCTTGTAAATCTAGATGATGCTTCTTTTATAATTC 2373
2280 TGTGGACACAGCTGATAGAGCTCTCTTGTAAATCTAGATGATGCTTCTTTTATAATTC 2339
2374 TACCTATTAGCTTTCTTTTATTTTAACTTTTGGGGGAAAAAGTATTTTGAAGT 2433
2340 TACCTATTAGCTTTCTTTTATTTTAACTTTTGGGGGAAAAAGTATTTTGAAGT 2399
2434 TGTCTTGAAGTATTTTAACTTTTAACTTTTGGGGGAAAAAGTATTTTGAAGT 2493
2400 TGTCTTGAAGTATTTTAACTTTTAACTTTTGGGGGAAAAAGTATTTTGAAGT 2459
2494 A 2494
2460 A 2460

```

```

RESULT 7
HUMGFBPBA 2474 bp mRNA linear PRI 08-NOV-1994
OCUS Human growth hormone-dependent insulin-like growth factor-binding
DEFINITION protein mRNA, complete cds.
ACCESSION M31159.1 GI:183115
VERSION M31159.1 GI:183115
KEYWORDS insulin-like growth factor binding protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2474)
AUTHORS Wood,W.I., Cachianes,G., Henzel,W.J., Winslow,G.A., Spencer,S.A.,
Heilmann,R., Martin,J.L. and Baxter,R.C.
TITLE Cloning and expression of the growth hormone-dependent insulin-like
growth factor-binding protein
JOURNAL Mol. Endocrinol. 2 (12), 1176-1185 (1988)
MEDLINE 89112197
PubMed 2464130
COMMENT Original source text: Human plasma, cdna to mRNA, clone BP-53.
FEATURES
Location/Qualifiers
1..2474
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="7p13-p12"
1..2474
/gene="IGFBP1"
<1..2474
/gene="IGFBP1"

```

```

/product="GFIBP mRNA"
110..985
/gene="IGFBP1"
/note="insulin-like growth factor-binding protein"
/codon_start=1
/protein_id="AA052541.1"
/db_xref="GI:183115"
/translation="MQRAPTLMAAALTLVLRGPPVVARAGASSGGLGPPVVRCEPD
ARALACAPPAVCAELVREPGCCGLTALCEGQPCGIYTERCGSLRCQPPDAR
PLQALLDGLCLCNASVRLRAYLLPAPPAGNASEBEDRAGSVSPSVSTHRV
SDKFEPLHSLKIIIIKKHAKDSQRYKVDYOSDTDFNSBSKRETEVPCRRRE
DTLHLKFLNVLSPRGVHIFPNCDDKGFYKKQCRPSKRGKRGFCWCVDYKQPLPGVT
TKKEDVHCYSMOSK"
ORIGIN 132 bp upstream of PfIMI site.
Query Match 97.4%; Score 2430.4; DB 9; Length 2474;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2464; Conservative 0; Mismatches 1; Indels 3; Gaps 3;
QY 12 CAGCGGCTTCTGCTGGATTCCACAGCTTTCGCGCGCTGCTGCTGCGCCCATCCCTCGG 71
DB 10 CAGCGGCTTCTGCTGGATTCCACAGCTTTCGCGCGCTGCTGCTGCGCCCATCCCTCGG 69
QY 72 CGCCACGCTGCCAAGCAGCGCTGCCCGGTTGAGGCGCTCATGACGCGCGCGGACCCAC 131
DB 70 CGCCACGCTGCCAAGCAGCGCTGCCCGGTTGAGGCGCTCATGACGCGCGCGGACCCAC 129
QY 132 GCTTGGGGCGCTGCGGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 191
DB 130 GCTTGGGGCGCTGCGGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189
QY 192 TGGCGGAGCTCGGGGGCTTGGTCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 251
DB 190 TGGCGGAGCTCGGGGGCTTGGTCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 249
QY 252 ACTGGCGAGTGGCGGCTTCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 311
DB 250 ACTGGCGAGTGGCGGCTTCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309
QY 312 CGGCTGCTGCTGAGTGGCGCTGAGCGAGGCGCAGCGCTGCGCGCTGCGCGCTGCGCGCTG 371
DB 310 CGGCTGCTGCTGAGTGGCGCTGAGCGAGGCGCAGCGCTGCGCGCTGCGCGCTGCGCGCTG 369
QY 372 CTGTGGTCCGGCTTCCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 431
DB 370 CTGTGGTCCGGCTTCCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 429
QY 432 GCTGGAGCGCGCGGCTTCCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 491
DB 430 GCTGGAGCGCGCGGCTTCCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 489
QY 492 GCTGGAGCGCGCGGCTTCCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 551
DB 490 GCTGGAGCGCGCGGCTTCCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 549
QY 552 CAGTGTGGAGCGCGGCTTCCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 611
DB 550 CAGTGTGGAGCGCGGCTTCCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 609
QY 612 CCTCCATTCAAAGATAATCATCATCAAGAAAGGCGATGCTAAAGACAGCGCGCGCTACAA 671
DB 610 CCTCCATTCAAAGATAATCATCATCAAGAAAGGCGATGCTAAAGACAGCGCGCGCTACAA 669
QY 672 AGTTGACTAGAGTCTCAGAGCAGATACCGAGACTTCTCTCCGAGTCCGAGCGGGA 731
DB 670 AGTTGACTAGAGTCTCAGAGCAGATACCGAGACTTCTCTCCGAGTCCGAGCGGGA 729
QY 732 GACAGATATGCTCCCTGCGGTAGAGAAATGGAAGACACACTGAATCACTGAAGTTTCT 791
DB 730 GACAGATATGCTCCCTGCGGTAGAGAAATGGAAGACACACTGAATCACTGAAGTTTCT 789
QY 792 CAATGTGCTGAGTCCCGAGGGGTGTACATTTCCCACTGTGACAGAGAGGGATTTTATAA 851

```

Db	1867	GAGGATAAGAGCTCTTTAAAGGCAAGGCTTTATTTTTCATCTCTCAATCTTTTGTCTCTCTCTTAG	192
Qy	1932	CACAAATGTTAAAAAGAAATAGTATATATACAGAAACAGGAAGGAGAAATGGCTTCTGCTGGGGAGC	1991
Db	1927	CACAAATGTTAAAAAGAAATAGTATATATACAGAAACAGGAAGGAGAAATGGCTTCTGCTGGGGAGC	1986
Qy	1992	CCATCCAGGACACTGGGAGCACATAGAGATTACCCCATGTTTCTGTCGAACCTTAGAGTCAATT	2051
Db	1987	CCATCCAGGACACTGGGAGCACATAGAGATTACCCCATGTTTCTGTCGAACCTTAGAGTCAATT	2046
Qy	2052	CTCATGCTTTTCTTTTATAATTCACATATATACAGAGAGATATGTTCTTGTAAACATT	2111
Db	2047	CTCATGCTTTTCTTTTATAATTCACATATATACAGAGAGATATGTTCTTGTAAACATT	2106
Qy	2112	GTATACAACATACGCCCAATATAGTAAGATCTATCTAGATATCTCTAGATGAAATGTT	2171
Db	2107	GTATACAACATAGCCCAATATAGTAAGATCTATCTAGATATCTCTAGATGAAATGTT	2166
Qy	2172	AGAGATGCTATTGTATACAACTGTGGCCATGACTGAGGAAAGAGGCTCAGCCCCAGAGAC	2231
Db	2167	AGAGATGCTATTGTATACAACTGTGGCCATGACTGAGGAAAGAGGCTCAGCCCCAGAGAC	2226
Qy	2232	TGGGCTGTCTCTCCGGAGGCGCAAAACCCAGAGGCTCTGGCAAAAGTCAGGCTCAGGGAGAC	2291
Db	2227	TGGGCTGTCTCTCCGGAGGCGCAAAACCCAGAGGCTCTGGCAAAAGTCAGGCTCAGGGAGAC	2286
Qy	2292	TCTGCCCTGTGCAGACCTCGGTGTGGACACACGCTGCATAGAGCTCTCTTGAAAAACAG	2351
Db	2287	TCTGCCCTGTGCAGACCTCGGTGTGGACACACGCTGCATAGAGCTCTCTTGAAAAACAG	2346
Qy	2352	AGGGGTCTCAAGACATTCGCTCCACCTATTAGCTTTTCTTTATTTTAACTTTTGGG	2411
Db	2347	AGGGGTCTCAAGACATTCGCTCCACCTATTAGCTTTTCTTTATTTTAACTTTTGGG	2406
Qy	2412	GGGAAAAGTATTTTTCAGAAAGTTTGTCTTGCATAGTATTTTAAATAGTAAATAAGTTT	2471
Db	2407	GGGAAAAGTATTTTTCAGAAAGTTTGTCTTGCATAGTATTTTAAATAGTAAATAAGTTT	2466
Qy	2472	TTACCAATT 2479	
Db	2467	TTACCAATT 2474	
RESULT 8			
AX713809			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			

[illegible]

1081	TTTATAATTCGTGTTGGTGAACATGATTTTTTTTTTAAACCAAAAGTTTAAAGAGAGGTTT	1140
1090	TTTATAATTCGTGTTGGTGAACATGATTTTTTTTTTAAACCAAAAGTTTAAAGAGAGGTTT	1109
1141	TTGAAATGCCATATGTTCTCTTTGAATGGTAAACTTGAGCATCTTTTACATTTCCAGTAGT	1200
1110	TTGAAATGCCATATGTTCTCTTTGAATGGTAAACTTGAGCATCTTTTACATTTCCAGTAGT	1169
1201	CAGCAAAAGAGCAGTTTGAATTTCTTTGGCTTCCTATCAAAATATTTCAGAGACTCGAGC	1260
1170	CAGCAAAAGAGCAGTTTGAATTTCTTTGGCTTCCTATCAAAATATTTCAGAGACTCGAGC	1229
1261	ACAGACCCAGACTTCATCGCGCCGTGGAAATGCTCACCACATGTTGGTCAAGAGCGGCCGA	1320
1230	CAGACCCAGACTTCATCGCGCCGTGGAAATGCTCACCACATGTTGGTCAAGAGCGGCCGA	1289
1321	CCAATGACTTTGTGACTTATAGCGGGCTGTGTGCCCTATGTAGAGAAACAGCTTCAACCCCA	1380
1290	CCAATGACTTTGTGACTTATAGCGGGCTGTGTGCCCTATGTAGAGAAACAGCTTCAACCCCA	1349
1381	CTCCCGGTACAGTGGCGACAGGCTTTATCGAGAAATAGAGAAACCTTTAAACCCGGGTCA	1440
1350	CTCCCGGTACAGTGGCGACAGGCTTTATCGAGAAATAGAGAAACCTTTAAACCCGGGTCA	1409
1441	CGGACATCCCAACGCAATGCTCTCTGGAGTCTACAGCCTTCTGTGGTGTCAATTTCTGAAAC	1500
1410	CGGACATCCCAACGCAATGCTCTCTGGAGTCTACAGCCTTCTGTGGTGTCAATTTCTGAAAC	1469
1501	AAGCGGTGGATCCTCTCAACCAAGAGAAATGTTATGTCTCAAGTGACCTGTACTGCTT	1560
1470	AAGCGGTGGATCCTCTCAACCAAGAGAAATGTTATGTCTCAAGTGACCTGTACTGCTT	1529
1561	GGGACTATTGGAGAAATAAGGTGGAGTCTCTACTTGTTTAAAAAATATGTATCAAGAA	1620
1530	GGGACTATTGGAGAAATAAGGTGGAGTCTCTACTTGTTTAAAAAATATGTATCAAGAA	1589
1621	TGTTCTAGGCACTCTGGGAACTTATAAAGCAGAGTATTTTGGGGCCTCTCTTCAGGAA	1680
1590	TGTTCTAGGCACTCTGGGAACTTATAAAGCAGAGTATTTTGGGGCCTCTCTTCAGGAA	1649
1681	TCTTCTCTAGACATGCCCAGTCTGAAGGCCCAGGATGGCTTTTGTCTGGGCCCTCGTGGG	1740
1650	TCTTCTCTAGACATGCCCCAGTCTGAAGGCCCAGGATGGCTTTTGTCTGGGCCCTCGTGGG	1709
1741	GTAGGAGGACAGAGACAGAGGAGTCTCAGCTCTCAATTCAGAGGATCAACAAGTAAT	1800
1710	GTAGGAGGACAGAGACAGAGGAGTCTCAGCTCTCAATTCAGAGGATCAACAAGTAAT	1769
1801	GGCACAATCTTCGGATGACTGACAGAAATAGTGTGTTCTAGTTCACAACTCAAGACGA	1860
1770	GGCACAATCTTCGGATGACTGCGAAAAATAGTGTGTTTGTAGTTCACAACTCAAGACGA	1829
1861	AGCTTAATTTCTGAGGATAAGCTCTTTAAAGGCAAGACTTTATTTCTCTCATCTTTTT	1920
1830	AGCTTAATTTCTGAGGATAAGCTCTTTAAAGGCAAGACTTTATTTCTCTCATCTTTTT	1889
1921	GTCTCTTTAGCACAATGTAAAAAGAAATAGTAATATCAGAACAGAGAGGAGGAATGGCT	1980
1890	GTCTCTTTAGCACAATGTAAAAAGAAATAGTAATATCAGAACAGAGAGGAGGAATGGCT	1949
1981	TGCTGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTGTGAAAC	2040
1950	TGCTGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTGTGAAAC	2009
2041	TTAGAGTCAATTCATGCTTTTCTTTAATTCACACATATATGACAGAGAGATATGTC	2100
2010	TTAGAGTCAATTCATGCTTTTCTTTAATTCACACATATATGACAGAGAGATATGTC	2069
2101	TTGTTAACTTGTATACAACTAGCCCCCAATATAGTAAGATCTATCTAGATAATCCTA	2160
2070	TTGTTAACTTGTATACAACTAGCCCCCAATATAGTAAGATCTATCTAGATAATCCTA	2129
2161	GATGAAATGTTAGATGCTATTTGATACAACTGTGGCCATGACTGAGGAATAGAGCTCA	2220

Db 2130 CATGAAATGTTAGAGATGCTATTTCATCAACATGTGGCCATGACTGAGGAAAGAGGTCTCA 2189  
Qy 2221 CGCCAGAGACTGGGCTGCTCTCCGAGGAGCCAAACCAAGAGGTCTGCAAGTCAAG 2280  
Db 2190 CGCCAGAGACTGGGCTGCTCTCCGAGGAGCCAAACCAAGAGGTCTGCAAGTCAAG 2249  
Qy 2281 CTCAGGAGACTGCTGCCCTGTCTGAGACCTCGGTGTGGACACACGCTGCAATAGAGTCTC 2340  
Db 2250 CTCAGGAGACTGCTGCCCTGTCTGAGACCTCGGTGTGGACACACGCTGCAATAGAGTCTC 2309  
Qy 2341 CTTGAAACACAGAGGGTCTCAAGACATTCGCTTACCTATTAGCTTTCTTTATTTTTT 2400  
Db 2310 CTTGAAACACAGAGGGTCTCAAGACATTCGCTTACCTATTAGCTTTCTTTATTTTTT 2369  
Qy 2401 AACTTTTGGGGGAAAAGTATTTTGAAGAGTTGTCTTGCATATGATTATAAATAGT 2460  
Db 2370 AACTTTTGGGGGAAAAGTATTTTGAAGAGTTGTCTTGCATATGATTATAAATAGT 2429  
Qy 2451 AAATAAAGTTTTTACCATT 2479  
Db 2430 AAATAAAGTTTTTACCATT 2448

AK055632 2448 bp mRNA linear PRI 01-AUG-2002  
Homo sapiens CDNA FLJ131070 fib, clone HSYRA2001332, highly similar  
to INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 3 PRECURSOR.  
AK055632  
AK055632.1 GI:16550409  
oligo capping; fib (full insert sequence).  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1  
Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,  
Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,  
Otsuki,T., Sato,H., Wakamatsu,A., Iehi,S., Yamamoto,J., Iseno,Y.,  
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,  
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,  
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,  
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,  
Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2448)  
Isogai,T., Otsuki,T. and Sugiyama,T.  
Direct Submission  
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: RAB and  
HRI.

FEATURES  
source  
1..2448  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HSYRA2001332"  
/cell\_type="synovocytes from rheumatoid arthritis  
(HS-RÅ)"  
/clone\_lib="HSYRA2"  
/note="cloning vector: pWE18FLJ3-primary culture,  
synovocytes from rheumatoid arthritis"

ORIGIN

Query Match 94.7%; Score 2361.8; DB 9; Length 2448;  
Best Local Similarity 97.8%; Pred. No. 0;  
Matches 2425; Conservative 0; Mismatches 2; Indels 52; Gaps 1;  
Qy 1 GGGCTGAGGATCAGCGGCTTCTGCTGGATTTCACAGCTTTCGGGCCGTGTACTGTGCGCC 60  
Db 22 GCGCTGAGGATCAGCGGCTTCTGCTGGATTTCACAGCTTTCGGGCCGTGTACTGTGCGCC 81  
Qy 61 CCATCCCTGCGGCCCGCCAGCTGCCAAGCAGGCTGCCCGGTTGCAGGCGTCAATGACAGCG 120  
Db 82 CCATCCCTGCGGCCCGCCAGCTGCCAAGCAGGCTGCCCGGTTGCAGGCGTCAATGACAGCG 141  
Qy 121 GCGGACCCACGCTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
Db 142 GCGGACCCACGCTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 201  
Qy 181 GTGGGCGGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 240  
Db 202 GTGGGCGGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 257  
Qy 241 GACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
Db 258 -----GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 269  
Qy 301 GAGCGGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
Db 270 GAGCGGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 329  
Qy 361 TACACGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
Db 330 TACACGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 389  
Qy 421 CTGCGAGCGCTCTGCTGAGCGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
Db 390 CTGCGAGCGCTCTGCTGAGCGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 449  
Qy 481 CGCGCTTACTCTGCTGAGCGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
Db 450 CGCGCTTACTCTGCTGAGCGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 509  
Qy 541 CGCAGCGCGCGCGCTGCTGAGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 510 CGCAGCGCGCGCGCTGCTGAGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 569  
Qy 601 AAGTTCACCCCTCTCATTCAAAAGATAATCATCATCAAGAAAGGCGTCTTAAACACAGC 660  
Db 570 AAGTTCACCCCTCTCATTCAAAAGATAATCATCATCAAGAAAGGCGTCTTAAACACAGC 629  
Qy 661 CAGCGCTACAAAGTTGACTTACGAGTCTCAGAGCAGATACCCAGAACTTCTCTCCCGAG 720  
Db 630 CAGCGCTACAAAGTTGACTTACGAGTCTCAGAGCAGATACCCAGAACTTCTCTCCCGAG 689  
Qy 721 TCAAGCGGAGACAGAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 690 TCAAGCGGAGACAGAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 749  
Qy 781 CTGAAGTTCCTCAATGTGCTGAGTCTCCAGGGTGTACATTCCTCCAACTGTGACAAAGAG 840  
Db 750 CTGAAGTTCCTCAATGTGCTGAGTCTCCAGGGTGTACATTCCTCCAACTGTGACAAAGAG 809  
Qy 841 GGATTTTATAAGAAAAGCAGTGTGCGCTTCAAAGGCGAGGAGCGGGGCTTCTGCTGG 900  
Db 810 GGATTTTATAAGAAAAGCAGTGTGCGCTTCAAAGGCGAGGAGCGGGGCTTCTGCTGG 869  
Qy 901 TGTGTGTAATGATGAGGCGAGCTCTCCAGGCTACCCAGGCTACCCAGGAGGAGGAGGAGCGTG 960  
Db 870 TGTGTGTAATGATGAGGCGAGCTCTCCAGGCTACCCAGGCTACCCAGGAGGAGGAGGAGCGTG 929  
Qy 961 CACTGTCTACAGCATGACAGAGCAAGTAGACGCTGCGCAAGGTTAATGTGAGAGTCAAT 1020  
Db 930 CACTGTCTACAGCATGACAGAGCAAGTAGACGCTGCGCAAGGTTAATGTGAGAGTCAAT 989

2y 1021 ATGCCTATTATTTGCACAAAAGATGCGCAAGGACATGACCAGCAGCTGGCTACAGCCTCGA 1080  
 Db 990 ATGCCTATTATTTGCACAAAAGATGCGCAAGGACATGACCAGCAGCTGGCTACAGCCTCGA 1049  
 2y 1081 TTTATATTTCTGTTTGGTGAAGTGAATTTTTTTTAAACCAAGTTTGAAGAGGTTT 1140  
 Db 1050 TTTATATTTCTGTTTGGTGAAGTGAATTTTTTTTAAACCAAGTTTGAAGAGGTTT 1109  
 2y 1141 TTGAATGCTATGTTTCTTTGAAATGTTAAACTTGAGCATCTTTTCACTTTCCAGTAGT 1200  
 Db 1110 TTGAATGCTATGTTTCTTTGAAATGTTAAACTTGAGCATCTTTTCACTTTCCAGTAGT 1169  
 2y 1201 CAGCAAGAGCAGTTGAATTTCTTCTGCTGCTTCCCTATCAAAATATTCAGAGACTCGAGC 1260  
 Db 1170 CAGCAAGAGCAGTTGAATTTCTTCTGCTGCTTCCCTATCAAAATATTCAGAGACTCGAGC 1229  
 2y 1261 ACAGCACCAGACTTCATGCGCCCGTGGAAATGCTCACCATGTTGCTGAAGCGGCCGA 1320  
 Db 1230 ACAGCACCAGACTTCATGCGCCCGTGGAAATGCTCACCATGTTGCTGAAGCGGCCGA 1289  
 2y 1321 CCAGTGAATTTGACTTAGCGGCTGTGTTGCTATGTAGAGAACGCTTCAACCCCA 1380  
 Db 1290 CCAGTGAATTTGACTTAGCGGCTGTGTTGCTATGTAGAGAACGCTTCAACCCCA 1349  
 2y 1381 CTCCTCCGTACAGTCGCGACAGGCTTTATCGAATAGGAAACCTTTAAACCCCGGTCTAT 1440  
 Db 1350 CTCCTCCGTACAGTCGCGACAGGCTTTATCGAATAGGAAACCTTTAAACCCCGGTCTAT 1409  
 2y 1441 CCGGACATCCCAAGCATGCTCTGGAGCTCAAGCCCTCTGTGGTGTCTATTTCTGAAC 1500  
 Db 1410 CCGGACATCCCAAGCATGCTCTGGAGCTCAAGCCCTCTGTGGTGTCTATTTCTGAAC 1469  
 2y 1501 AAGGGCTGGAATCCTCAACCAAGAAATGTTTATGCTTTCAAGTGACCTGACTGCTT 1560  
 Db 1470 AAGGGCTGGAATCCTCAACCAAGAAATGTTTATGCTTTCAAGTGACCTGACTGCTT 1529  
 2y 1561 GGGGACTATTGGAGAAATAGTGGAGTCTACTTGTTTTAAATAATATGATCTAAGAA 1620  
 Db 1530 GGGGACTATTGGAGAAATAGTGGAGTCTACTTGTTTTAAATAATATGATCTAAGAA 1589  
 2y 1621 TGTTCTAGGCACTCTGGGACCTTAAAGCAGGATTTTCGGGCTCTCTCTCAGGA 1680  
 Db 1590 TGTTCTAGGCACTCTGGGACCTTAAAGCAGGATTTTCGGGCTCTCTCTCAGGA 1649  
 2y 1681 TCTTCTGAACACATGCGCCAGTCTGAAGGCCAGGATGGCTTTTGTCTGGGCCCCGCTGGG 1740  
 Db 1650 TCTTCTGAACACATGCGCCAGTCTGAAGGCCAGGATGGCTTTTGTCTGGGCCCCGCTGGG 1709  
 2y 1741 GTAGAGGGACAGAGACAGGAGAGTCAAGCTCCACATTCAGAGCATCAAGATAT 1800  
 Db 1710 GTAGAGGGACAGAGACAGGAGAGTCAAGCTCCACATTCAGAGCATCAAGATAT 1769  
 2y 1801 GGCAAAATCTTCGGATGACCTGCAGAAATAGTGTGTTTGTAGTTTCAACAACTCAAGACGA 1860  
 Db 1770 GGCAAAATCTTCGGATGACCTGCAGAAATAGTGTGTTTGTAGTTTCAACAACTCAAGACGA 1829  
 2y 1861 AGCTTATTTCTGAGGATAAGCTTTTAAAGCAAGCTTTTATTTATCTCTCATCTTTT 1920  
 Db 1830 AGCTTATTTCTGAGGATAAGCTTTTAAAGCAAGCTTTTATTTATCTCTCATCTTTT 1889  
 2y 1921 GTCCTCTTAGCACAAATGTAAGAAATAGTAAATATCAGAACAGGAGAGGAAATGGCT 1980  
 Db 1890 GTCCTCTTAGCACAAATGTAAGAAATAGTAAATATCAGAACAGGAGGAGGAAATGGCT 1949  
 2y 1981 TGCTGGGAGCCCATCCAGGACATCGGAGACATAGAGATTCACCATGTTTGTGTAAC 2040  
 Db 1950 TGCTGGGAGCCCATCCAGGACATCGGAGACATAGAGATTCACCATGTTTGTGTAAC 2009  
 2y 2041 TTAGAGTATTTCTATGCTTTTCTTTTATATTTACATATATGACAGAGATATGTTT 2100  
 Db 2010 TTAGAGTATTTCTATGCTTTTCTTTTATATTTACATATATGACAGAGATATGTTT 2069  
 2y 2101 TTGTTAAATTTGTATACAAACATAGCCCAATATATAGTAAGATCTTATAGTAATCTCTA 2160

Db 2070 TTGTTAAATTTGTATACAAACATAGCCCAATATAGTAAGATCTTATAGTAATCTCTA 2129  
 Qy 2161 GATGAAATGTTTAGAGATGCTATTTGATACAACTGTGGCCATGACTGAGGAAAGAGCTCA 2220  
 Db 2130 GATGAAATGTTTAGAGATGCTATTTGATACAACTGTGGCCATGACTGAGGAAAGAGCTCA 2189  
 Qy 2221 CGCCAGAGACTGGGCTGCTCTCCGGAGGCCAAACCCAGAAAGTCTGCAAAAGTCAAG 2280  
 Db 2190 CGCCAGAGACTGGGCTGCTCTCCGGAGGCCAAACCCAGAAAGTCTGCAAAAGTCAAG 2249  
 Qy 2281 CTCAGGAGACTGCTCCCTGCTGCGAGACCTCGGTGTGGACACAGCTGCATAGAGCTCTC 2340  
 Db 2250 CTCAGGAGACTGCTCCCTGCTGCGAGACCTCGGTGTGGACACAGCTGCATAGAGCTCTC 2309  
 Qy 2341 CTTGAAAAACAGAGGGGTCTCAAGACATTTCTCCCTACCTATTAGCTTTTCTTTATTTT 2400  
 Db 2310 CTTGAAAAACAGAGGGGTCTCAAGACATTTCTCCCTACCTATTAGCTTTTCTTTATTTT 2369  
 Qy 2401 RACTTTTTGGGGGGAAGATTTTGGAGAGTTTGTCTGCAATGTATTTATTAATAGT 2460  
 Db 2370 RACTTTTTGGGGGGAAGATTTTGGAGAGTTTGTCTGCAATGTATTTATTAATAGT 2429  
 Qy 2461 AAATAAAGTTTTTACCAAT 2479  
 Db 2430 AAATAAAGTTTTTACCAAT 2448

RESULT 10  
 AK122873  
 LOCUS  
 DEFINITION  
 Homo sapiens cDNA FLJ16507 fis, clone HCHON2000364, highly similar to INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 3 PRECURSOR.  
 ACCESSION  
 AK122873  
 VERSION  
 AK122873.1 GI:34528071  
 KEYWORDS  
 oligo capping; fis (full insert sequence).  
 SOURCE  
 Homo sapiens  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1  
 Tanigami, A., Fujiwara, T., Shibahara, T., Goto, Y., Hirao, M., Ozaki, K., Shimizu, F., Wakase, H., Ono, T., Hishigaki, H., Watanabe, T., Ozaki, K., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kaneshiro, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Negahari, K., Masuho, Y., Nagai, K. and Isogai, T.  
 NEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 2430)  
 Isogai, T. and Yamamoto, J.  
 Direct Submission  
 Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan, Fax:81-438-52-3986) (E-mail:genomics@hri.co.jp, tel:81-438-52-3975, Fax:81-438-52-3986)  
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.  
 Location/Qualifiers  
 1..2430  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="HCHON2000364"  
 /cell\_type="chondrocytes (HC)"  
 /clone\_lib="HCHON2"

FEATURES  
 source



/note="cloning vector: pME18SFL3-primary culture, chondrocytes"

ORIGIN

Query Match				92.6%; Score 2309.4; DB 9; Length 2430;			
Best Local Similarity				96.9%; Pred. No. 0;			
Matches 2403; Conservative 0; Mismatches 6; Indels 70; Gaps 2;							
QY	1	GGCTGAGGATACGCGCTTCTGCTCGATTCACAGCTTCGCGCGTGTACTGTGCGC	60				
DB	22	GGCTGAGGATACGCGCTTCTGCTCGATTCACAGCTTCGCGCGTGTACTGTGCGC	81				
QY	61	COATCCCTGGCGGCCAGCTGCCAAGCAGCGTGCCTGGTTGAGCGCTCATCAGCGG	120				
DB	82	COATCCCTGGCGGCCAGCTGCCAAGCAGCGTGCCTGGTTGAGCGCTCATCAGCGG	141				
QY	121	GGCGACCCAGCTCTGGCGCGCTGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	180				
DB	142	GGCGACCCAGCTCTGGCGCGCTGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	201				
QY	181	GTGGCGCGGCTGGCGAGCTCGGGGGCTTGGTCCGCTGGCTGGCGAGCCGCTG	240				
DB	202	GTGGCGCGGCTGGCGAGCTCGGGGGCTTGGTCCGCTGGCTGGCGAGCCGCTG	249				
QY	241	GACGCGCTGCACTGGCCCAAGTGGCGCGCTCGCGCGCGCTGGCGGAGCTGGTGGC	300				
DB	250	-----CGC	252				
QY	301	GAGCGGGCTGGGCTGTGCTGCTGAGCTGGCACTGAGCGAGGGCGAGCGTGGGGATC	360				
DB	253	GAGCGGGCTGGGCTGTGCTGCTGAGCTGGCACTGAGCGAGGGCGAGCGTGGGGATC	312				
QY	361	TACACCGAGCTGTGGCTTCGGCTTCGCTGCCAGCGCTGGCGCGAGCGCGGACCG	420				
DB	313	TACACCGAGCTGTGGCTTCGGCTTCGCTGCCAGCGCTGGCGCGAGCGCGGACCG	372				
QY	421	CTGAGGCGCTGCTGGACGCGCGCGGCTGTGGCTCAACGCTAGTGGCTGAGCGCGCTG	480				
DB	373	CTGAGGCGCTGCTGGACGCGCGCGGCTGTGGCTCAACGCTAGTGGCTGAGCGCGCTG	432				
QY	481	CGCGCTACTGCTGCGAGCGCGCGAGCTCCAGGAAATGCTAGTGGTGGGAGAGAC	540				
DB	433	CGCGCTACTGCTGCGAGCGCGCGAGCTCCAGGAAATGCTAGTGGTGGGAGAGAC	492				
QY	541	CGCAGCGCGGAGTGTGGAGAGCGCGTCCGCTCCAGACGCGAGCGGCTGTGATCCC	600				
DB	493	CGCAGCGCGGAGTGTGGAGAGCGCGTCCGCTCCAGACGCGAGCGGCTGTGATCCC	552				
QY	601	AACTTCCACCCCTCCATTCAGGATATCATCATCAAGAGGGCATGCTAAAGACAGC	660				
DB	553	AACTTCCACCCCTCCATTCAGGATATCATCATCAAGAGGGCATGCTAAAGACAGC	612				
QY	661	CAGCGCTACAAAGTTGACTACAGCTCTCAGAGCAGATACCCAGAACTTCTCTCCGAG	720				
DB	613	CAGCGCTACAAAGTTGACTACAGCTCTCAGAGCAGATACCCAGAACTTCTCTCCGAG	672				
QY	721	TCCAGCGGAGACAGAAATGGTCCCTGCTGAGGAAATGGAGACACACTGAATCAC	780				
DB	673	TCCAGCGGAGACAGAAATGGTCCCTGCTGAGGAAATGGAGACACACTGAATCAC	732				
QY	781	CTGAAGTTCCTCAATGTGCTGAGTCCCGGGGTGTACACATTCCTCAACTGTGACAGAG	840				
DB	733	CTGAAGTTCCTCAATGTGCTGAGTCCCGGGGTGTACACATTCCTCAACTGTGACAGAG	792				
QY	841	GGATTTATAGAAAAGCAGTGTGCCCTTCCAGGCGAGGAGCGGCTTCTGCTGG	900				
DB	793	GGATTTATAGAAAAGCAGTGTGCCCTTCCAGGCGAGGAGCGGCTTCTGCTGG	852				
QY	901	TGTGTGATAAGTATGGGAGCGCTTCCCGAGGCTACACCAAGGGAAGAGGAGCGTG	960				
DB	853	TGTGTGATAAGTATGGGAGCGCTTCCCGAGGCTACACCAAGGGAAGAGGAGCGTG	912				
QY	961	CACTGCTACAGCATGCAGAGCAAGTAGAGCGCTCGCGCAAGGTTAATGTGGAGCTCAAT	1020				

DB	913	CACTGCTACAGCATGCAGAGCAAGTAGAGCGCTCGCGCAAGGTTAATGTGGAGCTCAAT	972				
QY	1021	ATGCTCTATTTTGCACAAAAGACTGCCAAGGACATGACAGCAGCTGGCTACAGCTCGA	1080				
DB	973	ATGCTCTATTTTGCACAAAAGACTGCCAAGGACATGACAGCAGCTGGCTACAGCTCGA	1032				
QY	1081	TTTATATTTCTGTTGTTGGTGAATGTTTAAACCAAGTTTAAAGAGGTTT	1140				
DB	1033	TTTATATTTCTGTTGTTGGTGAATGTTTAAACCAAGTTTAAAGAGGTTT	1091				
QY	1141	TTGAAATGCTATGTTTCTTTTGAATGTTAAACTTTCAGCATCTTTTCCAGTAGT	1200				
DB	1092	TTGAAATGCTATGTTTCTTTTGAATGTTAAACTTTCAGCATCTTTTCCAGTAGT	1151				
QY	1201	CAGCAAGAGCAGTTTGAATTTTCTGCTGCTTCTATCAAAATATTAGAGACTCGAGC	1260				
DB	1152	CAGCAAGAGCAGTTTGAATTTTCTGCTGCTTCTATCAAAATATTAGAGACTCGAGC	1211				
QY	1261	ACAGCAACCCAGACTTTCATGCGCGCGTGGAAATGCTCACACATGTTGGTCGAAGCGCGCA	1320				
DB	1212	ACAGCAACCCAGACTTTCATGCGCGCGTGGAAATGCTCACACATGTTGGTCGAAGCGCGCA	1271				
QY	1321	CACTGACTTTGTGACTTAGGCGGCTGTGCTGCTATGTAGAGAACAGCTTTCACCCCA	1380				
DB	1272	CACTGACTTTGTGACTTAGGCGGCTGTGCTGCTATGTAGAGAACAGCTTTCACCCCA	1331				
QY	1381	CTCCCGTACAGTGGCAGAGCTTTATCGAGAAATAGGAAACCTTTAAACCCCGTCAAT	1440				
DB	1332	CTCCCGTACAGTGGCAGAGCTTTATCGAGAAATAGGAAACCTTTAAACCCCGTCAAT	1391				
QY	1441	CGGACATCCCAACGATGCTCTGAGCTCAAGCTTCTGTGTGCTATTTCTGAAC	1500				
DB	1392	CGGACATCCCAACGATGCTCTGAGCTCAAGCTTCTGTGTGCTATTTCTGAAC	1451				
QY	1501	AAGGGCTGCTGCTCAACCAAGAGATGTTATCTTCAAGTACCTGTACTGCTT	1560				
DB	1452	AAGGGCTGCTGCTCAACCAAGAGATGTTATCTTCAAGTACCTGTACTGCTT	1511				
QY	1561	GGGACTATTGGAGAAATTAAGTGGAGTCTTCTGTTTAAATAATGTATCTAAGAA	1620				
DB	1512	GGGACTATTGGAGAAATTAAGTGGAGTCTTCTGTTTAAATAATGTATCTAAGAA	1571				
QY	1621	TGTTCTAGGCACTCTGGAACTTAAAGCAGTATTTTCGGGCCCTCTCTCAGGAA	1680				
DB	1572	TGTTCTAGGCACTCTGGAACTTAAAGCAGTATTTTCGGGCCCTCTCTCAGGAA	1631				
QY	1681	TCCTCTGAAGCATGCGCCAGTCCGAAGGCCAGGATGCTTTCGTCGGGCCCTGGG	1740				
DB	1632	TCCTCTGAAGCATGCGCCAGTCCGAAGGCCAGGATGCTTTCGTCGGGCCCTGGG	1691				
QY	1741	GTAGGAGGACAGAGACAGGAGAGTCCAGCTCCACATTCAGAGGCATCAAGTAAT	1800				
DB	1692	GTAGGAGGACAGAGACAGGAGAGTCCAGCTCCACATTCAGAGGCATCAAGTAAT	1751				
QY	1801	GGCACAATTTCTCGGATGACTGCAGAAAAATAGTGTCTTGTAGTTCAAACACTCAAGACA	1860				
DB	1752	GGCACAATTTCTCGGATGACTGCAGAAAAATAGTGTCTTGTAGTTCAAACACTCAAGACA	1811				
QY	1861	AGCTTATTTCTGAGGATAAGCTTTTAAAGGCAAGCTTTATTTTCATCTCTCATCTTTT	1920				
DB	1812	AGCTTATTTCTGAGGATAAGCTTTTAAAGGCAAGCTTTATTTTCATCTCTCATCTTTT	1871				
QY	1921	GTCTCTCTTAGCAATGTAAAAAGAAATAGTAATATCAGAAACAGGAAGGAGATGGCT	1980				
DB	1872	GTCTCTCTTAGCAATGTAAAAAGAAATAGTAATATCAGAAACAGGAAGGAGATGGCT	1931				
QY	1981	TGCTGGGAGCCCATCCAGGACACTGGGAGCACAAGAGATTCAACCATTTGTTTGAAC	2040				
DB	1932	TGCTGGGAGCCCATCCAGGACACTGGGAGCACAAGAGATTCAACCATTTGTTTGAAC	1991				
QY	2041	TTAGAGTCACTTCTCATGCTTTTCTTTTAAATTCACATATATGCAGAGAAATATGTTTC	2100				



```

2b 1992 TTAGAGTCATCTCATGCTTTCTTTATTAATTCACATATATATGCGAGAGATATGTTTC 2051
2y 2101 TTGTTAACTTGTATACAAATAGCCCAATATAGTAAGATCTATACATAGATAATCTTA 2160
2b 2052 TTGTTAACTTGTATACAAATAGCCCAATATAGTAAGATCTATACATAGATAATCTTA 2111
2y 2161 GATGAATGTATAGATGCTATTTGATACAACTGTGGCCATGACTGAGGAAAGGAGCTCA 2220
2b 2112 GATGAATGTATAGATGCTATTTGATACAACTGTGGCCATGACTGAGGAAAGGAGCTCA 2171
2y 2221 CGCCGAGAGACTGGGCTGCTCTCCCGAGGCCAAACCCCAAGAGGTCTGGCAAGTTCAGG 2280
2b 2172 CGCCGAGAGACTGGGCTGCTCTCCCGAGGCCAAACCCCAAGAGGTCTGGCAAGTTCAGG 2231
2y 2281 CTGAGGAGACTCTGCCCTGCTCAGACCTCGGTGTGGACACACGCTGATAGAGCTCTC 2340
2b 2232 CTCAGGAGACTCTGCCCTGCTCAGACCTCGGTGTGGACACACGCTGATAGAGCTCTC 2291
2y 2341 CTTGAAACACAGAGGGGTCTCAAGACATCTCGCTACTATAGCTTTCTTTATTTTTT 2400
2b 2292 CTTGAAACACAGAGGGGTCTCAAGACATCTCGCTACTATAGCTTTCTTTATTTTTT 2351
2y 2401 AACTTTTGGGGGAAAGTATTTTGGAGAGTTTGTCTTGCAATGATTTATAATAGT 2460
2b 2352 AACTTTTGGGGGAAAGTATTTTGGAGAGTTTGTCTTGCAATGATTTATAATAGT 2411
2y 2461 AATAAGTTTTTACCATT 2479
2b 2412 AATAAGTTTTTACCATT 2430

RESULT 11
AK075483
LOCUS AK075483 2412 bp mRNA linear PRI 03-SEP-2002
DEFINITION Homo sapiens cDNA PSE00177 f1s, clone OVARC1000287, highly similar
to Human growth hormone-dependent insulin-like growth
factor-binding protein mRNA.
ACCESSION AK075483
VERSION 1.1 GI:22761657
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Ota, T., Nishikawa, T., Suzuki, Y., Kawai-Hio, Y., Hayashi, K.,
Ishii, S., Saito, K., Yamamoto, J., Wakamatsu, A., Nagai, T.,
Nakamura, Y., Sato, H., Nagahara, K., Sugano, S. and Isogai, T.
HRI human cDNA sequencing project
Unpublished
2 (bases 1 to 2412)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (20-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
HRI human cDNA sequencing project; cDNA 5' & 3'-end one pass
sequencing, clone selection and full insert sequencing; Helix
Research Institute (supported by Japan Key Technology Center etc.);
cDNA library construction; Institute of Medical Science, University
of Tokyo, Laboratory of Genome Structure, Human Genome Center.
FEATURES
Location/Qualifiers
1..2412
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OVARC1000287"
/tissue_type="ovary, tumor tissue"
/clone_lib="OVARC1"
/note="Cloning vector: pME18SFL3"
239..240
misc_feature
/note="86 bases segment is present in M31159 and X64875,
probably caused by alternative splicing"

```

```

ORIGIN
Query Match 91.6%; Score 2283.4; DB 9; Length 2412;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 2391; Conservative 0; Mismatches 1; Indels 87; Gaps 2;

QY 1 CGCGTAGGATCAGCCGCTTCCTGCTGGATTCCACAGCTTCGCGCGTGTACTGTGCGCC 60
DB 21 CGCGTAGGATCAGCCGCTTCCTGCTGGATTCCACAGCTTCGCGCGTGTACTGTGCGCC 80
QY 61 CCATCCCTGCGCGCCCAAGCCTGCCAAGCAGCGTCCCGGTTCAGCGCTGATCATCAGCG 120
DB 81 CCATCCCTGCGCGCCCAAGCCTGCCAAGCAGCGTCCCGGTTCAGCGCTGATCATCAGCG 140
QY 121 CGCGAGCCAGCTCTGGCGCGCTGCGCTGACTGCTGTGCTGTGCTCCGCGGCGCGCG 180
DB 141 CGCGAGCCAGCTCTGGCGCGCTGCGCTGACTGCTGTGCTGTGCTCCGCGGCGCGCG 200
QY 181 GTGCGCGCGGCTGCGCGGAGCTCGGGGGGCTTGGGTCCCGTGTGCTGCGCTGCGAGCC 240
DB 201 GTGCGCGCGGCTGCGCGGAGCTCGGGGGGCTTGGGTCCCGTGTGCTGCGAGCTGCTG 237
QY 241 GAGCGCGTGCACCTGCGCGCAGTGCAGCGCTCCGCGCGCTGTGCGCGAGCTGCTGCGCG 300
DB 238 ----- 237
QY 301 GAGCGCGCTGCGGCTGCTGCTGACGTGCGCACTGAGCGAGGCGCCAGCCGTGCGGCATC 360
DB 238 ---CGGGCTGCGGCTGCTGCTGACGTGCGCACTGAGCGAGGCGCCAGCCGTGCGGCATC 294
QY 361 TACACCGAGCGCTGTGCTCGGCTTCGCGCTTCGCTGCGAGCGCTGCGCGCGAGCGCGACCG 420
DB 295 TACACCGAGCGCTGTGCTCGGCTTCGCGCTTCGCTGCGAGCGCTGCGCGCGAGCGCGACCG 354
QY 421 CTGCAAGCGCTGCTGCGAGCGCGCGGCTCTGCGTCAACGCTAGTGTGCTGCGCGCGCTG 480
DB 355 CTGCAAGCGCTGCTGCGAGCGCGCGGCTCTGCGTCAACGCTAGTGTGCTGCGCGCGCTG 414
QY 481 CGCGCTTACCTGCTGCGAGCGCGCGCTCCAGGAAATGCTAGTGTGCGAGGAGAGAC 540
DB 415 CGCGCTTACCTGCTGCGAGCGCGCGCTCCAGGAAATGCTAGTGTGCGAGGAGAGAC 474
QY 541 CGCAGCGCGCGCAGTGTGAGAGAGCGCGTCCGCTCCAGCAGCGCACCGGCTGTCTGATCCC 600
DB 475 CGCAGCGCGCGCAGTGTGAGAGAGCGCGTCCGCTCCAGCAGCGCACCGGCTGTCTGATCCC 534
QY 601 AAGTTCCACCCCTCCATTTCAAAGATAATCATCATCAAGAAAGGCGCATGTCTGATGAGCAGC 660
DB 535 AAGTTCCACCCCTCCATTTCAAAGATAATCATCATCAAGAAAGGCGCATGTCTGATGAGCAGC 594
QY 661 CAGCGCTACAAGTTGACTACGAGTCTCAGAGCAGCATACCCAGAACTTCTCTCTCGAG 720
DB 595 CAGCGCTACAAGTTGACTACGAGTCTCAGAGCAGCATACCCAGAACTTCTCTCTCGAG 654
QY 721 TCCAGCGCGGAGACAGAAATATGTCCTCCGCTAGAGAAATGGAAGACACACTGAATCAC 780
DB 655 TCCAGCGCGGAGACAGAAATATGTCCTCCGCTAGAGAAATGGAAGACACACTGAATCAC 714
QY 781 GTGAAGTTCTCAATGCTGAGTCCCGGGGTGTACATTTCCCACTGTGCAAGAG 840
DB 715 GTGAAGTTCTCAATGCTGAGTCCCGGGGTGTACATTTCCCACTGTGCAAGAG 774
QY 841 GGATTTTATAAGAAAAAGCAGTGTCCGCTTCCAAAGCAGAGAGCGGGCTTCTGTGG 900
DB 775 GGATTTTATAAGAAAAAGCAGTGTCCGCTTCCAAAGCAGAGAGCGGGCTTCTGTGG 834
QY 901 TGTGTGGATAGATATGGGCGAGCTCTTCCAGGCTACACCACCAAGGGGAGGAGAGCTG 960
DB 835 TGTGTGGATAGATATGGGCGAGCTCTTCCAGGCTACACCACCAAGGGGAGGAGAGCTG 894
QY 961 CACTGCTACAGCATGCGAGCAAGTAGACGCTGCGCAAGGTTAATGTGAGGTCAAAAT 1020
DB 895 CACTGCTACAGCATGCGAGCAAGTAGACGCTGCGCAAGGTTAATGTGAGGTCAAAAT 954

```

QY 1021 ATGCTTATTTTGCACAAAAGACTGCGCAAGGACATGACACAGAGCTGGCTACAGGCTCGA 1080  
DB 955 ATGCTTATTTTGCACAAAAGACTGCGCAAGGACATGACACAGAGCTGGCTACAGGCTCGA 1014  
QY 1081 TTTATATTTCTGTTTGTGGTGAAGTCTGATTTTATTTTAAACCAAGTTTGAAGAAGGTTT 1140  
DB 1015 TTTATATTTCTGTTTGTGGTGAAGTCTGATTTTATTTTAAACCAAGTTTGAAGAAGGTTT 1073  
QY 1141 TTGAATGCCATGCTTATTTCTTTTGAATGTTAACTTTGAGCATCTTTTCACTTTCCAGTAGT 1200  
DB 1074 TTGAATGCCATGCTTATTTCTTTTGAATGTTAACTTTGAGCATCTTTTCACTTTCCAGTAGT 1133  
QY 1201 CAGGAAGACAGCTTTGAATTTCTTGTGCTTCTTATCCTATCAAAATATTCAGACATCGAGC 1260  
DB 1134 CAGGAAGACAGCTTTGAATTTCTTGTGCTTCTTATCCTATCAAAATATTCAGACATCGAGC 1193  
QY 1261 ACAGCACCCAGACTTCAATGCGCGCTGGAAATGCTCACCATGTTGGTGGAGCGGCGA 1320  
DB 1194 ACAGCACCCAGACTTCAATGCGCGCTGGAAATGCTCACCATGTTGGTGGAGCGGCGA 1253  
QY 1321 CCACTGACTTTGTGACTTTAGCGGCTGTGTGCTATGTAGAGAACACGCTTCAACCCCA 1380  
DB 1254 CCACTGACTTTGTGACTTTAGCGGCTGTGTGCTATGTAGAGAACACGCTTCAACCCCA 1313  
QY 1381 CTCCCGTACAGTGGCGACAGGCTTTATCGAGAAATAGGAAACCTTTAAACCCCGTCAAT 1440  
DB 1314 CTCCCGTACAGTGGCGACAGGCTTTATCGAGAAATAGGAAACCTTTAAACCCCGTCAAT 1373  
QY 1441 CCGGACATCCCAACGCAATGCTCTGAGGCTCACAGCTTCTGTGGTGTCAATTTGAAAC 1500  
DB 1374 CCGGACATCCCAACGCAATGCTCTGAGGCTCACAGCTTCTGTGGTGTCAATTTGAAAC 1433  
QY 1501 AAGGCGTGTGATCCCTCAACCAAGACAGTATTTATGCTTCAAGTGACCTGTACTGCTT 1560  
DB 1434 AAGGCGTGTGATCCCTCAACCAAGACAGTATTTATGCTTCAAGTGACCTGTACTGCTT 1493  
QY 1561 GGGGACTATTTGGGAAATTAAGTGGAGTCTACTTCTTTTAAATAATATGATCAAGAA 1620  
DB 1494 GGGGACTATTTGGGAAATTAAGTGGAGTCTACTTCTTTTAAATAATATGATCAAGAA 1553  
QY 1621 TGTTCTAGGCACTCTGGGAACCTATTAAGGCGAGTATTTTCGGGCTCTCTTTCAGGAA 1680  
DB 1554 TGTTCTAGGCACTCTGGGAACCTATTAAGGCGAGTATTTTCGGGCTCTCTTTCAGGAA 1613  
QY 1681 TCTTCTCTGAAGACATGCCCCAGTCCGAAGGCCAGGATGGCTTTTGTGCGGCGCCGCTGGG 1740  
DB 1614 TCTTCTCTGAAGACATGCCCCAGTCCGAAGGCCAGGATGGCTTTTGTGCGGCGCCGCTGGG 1673  
QY 1741 GTAGGAGGACAGAGACAGGAGAGTCAAGCTCCACATTCAGAGGATCAAGTAAT 1800  
DB 1674 GTAGGAGGACAGAGACAGGAGAGTCAAGCTCCACATTCAGAGGATCAAGTAAT 1733  
QY 1801 GGCAATTTCTCGATGACTCGAGAAATAGTGTGTTTGTAGTTTCAACCACTCAAGACGA 1860  
DB 1734 GGCAATTTCTCGATGACTCGAGAAATAGTGTGTTTGTAGTTTCAACCACTCAAGACGA 1793  
QY 1861 AGCTTATTTCTGAGTAAGCTTTTAAAGCAAGCTTTATTTTCACTCTCACTTTT 1920  
DB 1794 AGCTTATTTCTGAGTAAGCTTTTAAAGCAAGCTTTATTTTCACTCTCACTTTT 1853  
QY 1921 GTCTCTCTTAGCAATGTATTAAGAAATAGTATATCAGAACAGGAGGAAATGGCT 1980  
DB 1854 GTCTCTCTTAGCAATGTATTAAGAAATAGTATATCAGAACAGGAGGAAATGGCT 1913  
QY 1981 TGCTGGGAGCCCATCCAGGACACTGGAGACCATAGAGATTCACCCATGTTTGTGAAC 2040  
DB 1914 TGCTGGGAGCCCATCCAGGACACTGGAGACCATAGAGATTCACCCATGTTTGTGAAC 1973  
QY 2041 TTAGAGTCATTTCTATGCTTTTCTTTATATTCACATATATGACAGAAATATGTTTC 2100  
DB 1974 TTAGAGTCATTTCTATGCTTTTCTTTATATTCACATATATGACAGAAATATGTTTC 2033

QY 2101 TTGTTAACTTTGTATACAAACATAGCCCAATATAGTAGATCTATCTAGATAATCCTA 2160  
DB 2034 TTGTTAACTTTGTATACAAACATAGCCCAATATAGTAGATCTATCTAGATAATCCTA 2093  
QY 2161 GATCAAAATGTTAGAGTGTCTATTTGATACAACTGTGGCCATGCTGAGGAAGGAGCTCA 2220  
DB 2094 GATCAAAATGTTAGAGTGTCTATATGATACAACTGTGGCCATGCTGAGGAAGGAGCTCA 2153  
QY 2221 CGCCACAGAGACTGGGCTGTCTCTCCGAGGCGCCAAACCCAAAGAGGCTGCGCAAAGTCAGG 2280  
DB 2154 CGCCACAGAGACTGGGCTGTCTCTCCGAGGCGCCAAACCCAAAGAGGCTGCGCAAAGTCAGG 2213  
QY 2281 CTCAGGAGACTCTGCCCTGTGCTGAGACCTCGGTGGGACACACGCTGCAATAGAGCTCTC 2340  
DB 2214 CTCAGGAGACTCTGCCCTGTGCTGAGACCTCGGTGGGACACACGCTGCAATAGAGCTCTC 2273  
QY 2341 CTTGAAACAGAGGGGCTCTCAAGACATTTGCTCTACCTATTTAGCTTTCTTTATTTTTT 2400  
DB 2274 CTTGAAACAGAGGGGCTCTCAAGACATTTGCTCTACCTATTTAGCTTTCTTTATTTTTT 2333  
QY 2401 AACTTTTTGGGGGAAAAGTATTTTGGAGAGTTTGTCTTGCATGTATTTATAAATAGT 2460  
DB 2334 AACTTTTTGGGGGAAAAGTATTTTGGAGAGTTTGTCTTGCATGTATTTATAAATAGT 2393  
QY 2461 AATAAAGTTTTTACCATT 2479  
DB 2394 AATAAAGTTTTTACCATT 2412

RESULT 12  
AK095907  
LOCUS Homo sapiens cDNA FLJ38588 fis, clone HCHON2009795.  
DEFINITION AK095907 mRNA linear  
ACCESSION AK095907  
VERSION AK095907.1 GI:21755258  
KEYWORDS cDNA capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,  
Arita, M., Mueashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,  
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,  
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,  
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,  
Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,  
Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and  
Isogai, T.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2364)  
AUTHORS Isogai, T. and Yamamoto, J.  
TITLE Direct Submission  
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing;  
Research Association for Biotechnology (RAB); cDNA library  
construction; Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing; HRI and  
RAB; annotation: HRI and RAB.  
FEATURES  
source  
1. 2364  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HCHON2009795"  
/cell\_type="chondrocytes (HC)"  
/clone\_lib="HCHON2"

/notes="cloning vector: pME18SFL3-primary culture, chondrocytes"

ORIGIN

Query Match 90.2%; Score 2248.6; DB 9; Length 2364;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 2273; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
 201 CTGGGGGGCTTGGGTCCCGTGGTGGCTGCGCTCGAGAGCGGTGGAGCCGCGTGCACCTGSCCCA 260  
 88 CTCGGCGGCTTGGGTCCCGTGGTGGCTGCGAGCGGTGGAGCCGCGTGCACCTGSCCCA 147  
 261 GTGGCGCCCTCCGCGCCGCTGTGCGCGAGCTGTGCGAGCCGCGGTGCGCGTGTG 320  
 148 GTGGCGCCCTCCGCGCCGCTGTGCGCGAGCTGTGCGAGCCGCGGTGCGCGTGTG 207  
 321 CCTGAAGCTGCGCACTGAGCGAGGCCAGCGGTGCGGCACTACACGAGCGGTGTGGCTC 380  
 208 CCTGAAGCTGCGCACTGAGCGAGGCCAGCGGTGCGGCACTACACGAGCGGTGTGGCTC 267  
 381 CGGCTTTCGCTGCGAGCGGTGCGCGAGCTGTGCGAGCCGCGGTGCGCGTGTG 440  
 268 CGGCTTTCGCTGCGAGCGGTGCGCGAGCTGTGCGAGCCGCGGTGCGCGTGTG 327  
 441 CGGCGGCTCTGCGTCAAGCTAGTCCGCTGAGCGCGCTTCCGCGCTTACCTGCTCCAGC 500  
 328 CGGCGGCTCTGCGTCAAGCTAGTCCGCTGAGCGCGCTTCCGCGCTTACCTGCTCCAGC 387  
 501 GCGCGAGCTCCAGGAATGCTAGTGTGCGAGGAGAGCCGAGCGCGGTGCTGCGA 560  
 388 GCGCGAGCTCCAGGAATGCTAGTGTGCGAGGAGAGCCGAGCGCGGTGCTGCGA 447  
 561 GAGCGCGCTCGCTCCAGCGAGCAGCGCGGTGCTGATCCCAAGTTCACCGCCCTCCATTC 620  
 448 GAGCGCGCTCGCTCCAGCGAGCAGCGCGGTGCTGATCCCAAGTTCACCGCCCTCCATTC 507  
 621 AAGATATCATCATCAAGAAAGGCGATGCTAAGACAGCGCGGTGCTAAGATGACTA 680  
 508 AAGATATCATCATCAAGAAAGGCGATGCTAAGACAGCGCGGTGCTAAGATGACTA 567  
 681 CGAGTCTCAGAGCAGATCCAGAACTTCTCTCCGAGTCCAAAGCGGAGACAGAATA 740  
 568 CGAGTCTCAGAGCAGATCCAGAACTTCTCTCCGAGTCCAAAGCGGAGACAGNATA 627  
 741 TGGTCCCTGCGTGAAGAAATGGAAGACACACTGAATCACTGAAGTTCCTCAATGTGCT 800  
 628 TGGTCCCTGCGTGAAGAAATGGAAGACACACTGAATCACTGAAGTTCCTCAATGTGCT 687  
 801 GAGTCCAGGGGTGACACATCCCACTGTGACAGAGGGATTTTATAGAAAGACGA 860  
 688 GAGTCCAGGGGTGACACATCCCACTGTGACAGAGGGATTTTATAGAAAGACGA 747  
 861 GTGTGCGCCCTTCCAAAGGACAGAGCGGGCTTCTGCTGCTGTGATAGATGGGCA 920  
 748 GTGTGCGCCCTTCCAAAGGACAGAGCGGGCTTCTGCTGCTGTGATAGATGGGCA 807  
 921 GCCTTCCAGGCTACACCAAGGGAGAGGAGCGTGCACCTGCTACAGCATGACAG 980  
 808 GCCTTCCAGGCTACACCAAGGGAGAGGAGCGTGCACCTGCTACAGCATGACAG 867  
 981 CAAGTAGAGCGCTGCGCGCAAGGTAAATGTGAGGTCAAAATATGCTTTATTTGACAAA 1040  
 868 CAAGTAGAGCGCTGCGCGCAAGGTAAATGTGAGGTCAAAATATGCTTTATTTGACAAA 927  
 1041 GACTGCCAAGACATGACAGAGCTGCTACAGCTCGATTTATTTCTGTTGTTGGT 1100  
 928 GACTGCCAAGACATGACAGAGCTGCTACAGCTCGATTTATTTCT- TTTGTTGGT 986  
 1101 GAACTGATTTTTTAAACCAAGTTTAAAGAGGTTTTTGAATGCTTATGCTTTCT 1160  
 987 GAACTGA- TTTTAAACCAAGTTTAAAGAGGTTTTTGAATGCTTATGCTTTCT 1045  
 1161 TTGAATGGTAAACTTGAGCATCTTTTCACTTCCAGTAGTCAGCAAGAGCAGTTTGAAT 1220

1046 TTGAATGGTAAACTTGAGCATCTTTTCACTTCCAGTAGTCAGCAAGAGCAGTTTGAAT 1105  
 1221 TTTCTTGTGCTTCTCTATCAAAATATTAGAGACTCGAGACAGCACCAGACTTCATGC 1280  
 1106 TTTCTTGTGCTTCTCTATCAAAATATTAGAGACTCGAGACAGCACCAGACTTCATGC 1165  
 1281 GCGCGTGAATGCTCACACATGTTGTCGAAGCGCGGACCACTGACTTTTGTGACTTAG 1340  
 1166 GCGCGTGAATGCTCACACATGTTGTCGAAGCGCGGACCACTGACTTTTGTGACTTAG 1225  
 1341 GCGCGTGTGTTGCTCTATGTCAGAGAACACGCTTACCCCACTCCCGGTACAGTGGGACA 1400  
 1226 GCGCGTGTGTTGCTCTATGTCAGAGAACACGCTTACCCCACTCCCGGTACAGTGGGACA 1285  
 1401 GCGTTTATCGAGATAGGAAACCTTTAAACCCCGGTATCCGAGACATCCCAACGCAATGC 1460  
 1286 GCGTTTATCGAGATAGGAAACCTTTAAACCCCGGTATCCGAGACATCCCAACGCAATGC 1345  
 1461 TCCTGGAGCTCACAGCCTTCTGTGTGTCTATTCTGAAACAAAGGCGGTGGATCCCTCAAC 1520  
 1346 TCCTGGAGCTCACAGCCTTCTGTGTGTCTATTCTGAAACAAAGGCGGTGGATCCCTCAAC 1405  
 1521 CAAGAAGATGTTTATGCTTCAAGTGACCTGTACTGCTTGGGACTATTGGAGAAATA 1580  
 1406 CAAGAAGATGTTTATGCTTCAAGTGACCTGTACTGCTTGGGACTATTGGAGAAATA 1465  
 1581 AGTGGAGTCTTACTTGTTTAAATAATATGATCTAAGAATGTTCTAGGGCACTCTGGGA 1640  
 1466 AGTGGAGTCTTACTTGTTTAAATAATATGATCTAAGAATGTTCTAGGGCACTCTGGGA 1525  
 1641 ACCTATAAAGGACGATATTTCGGGCCCTCTCTTCAGGAATCTTCTGAGACATGGCCC 1700  
 1526 ACCTATAAAGGACGATATTTCGGGCCCTCTCTTCAGGAATCTTCTGAGACATGGCCC 1585  
 1701 AGTGGAAAGCGCCAGGATGGCTTTTGTCTGCGGCCCGGTAGGAGGAGACAGAGAGACA 1760  
 1586 AGTGGAAAGCGCCAGGATGGCTTTTGTCTGCGGCCCGGTAGGAGGAGACAGAGAGACA 1645  
 1761 GGGAGATCAGCCTCCACATTCAGAGGATCAAGATTAATGGCAATTTCTCGATGAC 1820  
 1646 GGGAGATCAGCCTCCACATTCAGAGGATCAAGATTAATGGCAATTTCTCGATGAC 1705  
 1821 TGCAGAAATAGTGTGTTGTAGTTCAACAACTCAAGACGAAGCTTATTCTGAGGATAAG 1880  
 1706 TGCAGAAATAGTGTGTTGTAGTTCAACAACTCAAGACGAAGCTTATTCTGAGGATAAG 1765  
 1881 CTCTTTAAAGGCAAGCTTTATTTTCACTCTCATCTTTTGTCTCTTTAGCAATGTA 1940  
 1766 CTCTTTAAAGGCAAGCTTTATTTTCACTCTCATCTTTTGTCTCTTTAGCAATGTA 1825  
 1941 AAAAGATAGTAAATATCAGACAGGAGGAGGATGGCTTGTGGGAGGCCATCCAGG 2000  
 1826 AAAAGATAGTAAATATCAGACAGGAGGAGGATGGCTTGTGGGAGGCCATCCAGG 1885  
 2001 ACACCTGGGAGACATAGAGATTTCACCCATGTTTGTGAACCTTAGAGTCAATCTCATGCTT 2060  
 1886 ACACCTGGGAGACATAGAGATTTCACCCATGTTTGTGAACCTTAGAGTCAATCTCATGCTT 1945  
 2061 TTTCTTTAATTTACACATATATGAGAGAGATATGTTCTTTGTTAAACATTTGTATACAC 2120  
 1946 TTTCTTTAATTTACACATATATGAGAGAGATATGTTCTTTGTTAAACATTTGTATACAC 2005  
 2121 ATAGCCCCAAATATAGTAAATCTATAGATAATCTTAGATGAATGTTTAGAGATGCT 2180  
 2006 ATAGCCCCAAATATAGTAAATCTATAGATAATCTTAGATGAATGTTTAGAGATGCT 2065  
 2181 ATTGTATACACTGTGGCCATGACTGAGAAAGAGAGCTCACGCCAGAGACTGGGCTGCT 2240  
 2066 ATATGATACAACTGTGGCCATGACTGAGAAAGAGAGCTCACGCCAGAGACTGGGCTGCT 2125  
 2241 CTCGGGAGGCGCAACCCCAAGAGGCTTGGCAAGAGTCCAGGCTCAGGAGACTCTGCCCTG 2300

	D	B	C T C C G G A G C C C A A C C C A A G A G G T G G G C A A A G T C A G G C T C A G G G A G A C T C T G C C C T G 2185
	Q	y	C T C G A C A C C T C G G T G G G A C A C A C G C T C A T A G A G C T C T C C T T G A A A A C A G A G G G G T C T C 2360
	D	b	C T C G A C A C C T C G G T G G G A C A C A C G C T C A T A G A G C T C T C C T T G A A A A C A G A G G G G T C T C 2245
	Q	y	A G A C A T T C G C T A C C T A T F A G C T T T C T T A T T T T T T A A C T T T T T T G G G G G A B R A G T 2420
	D	b	A G A C A T T C G C T A C C T A T F A G C T T T C T T A T T T T T T A A C T T T T T T A A C T T T T T T T G G G G G A A A G T 2305
	Q	y	A T T T T T T G A G A A G T T T G C T G T C G A A T G T A T T T A T A A T A G T A A A A A A G T T T T T A C C A T T 2479
	D	b	A T T T T T G A G A A G T T T G C T G T G C A A T G T A T T T A T A A T A G T A T T A A A A G T T T T T A C C A T T 2364
RESULT 13			
AKO95869			
LOCUS			Homo sapiens cDNA FLJ38550 fis, clone HCHONZ002123, highly similar to INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 3 PRECURSOR.
DEFINITION			
ACCESSION			AKO95869 linear PRI 15-JUL-2002
VERSION			AKO95869.1 GI:21755214
KEYWORDS			Oligo capping; fis (full insert sequence).
SOURCE			Homo sapiens (human)
ORGANISM			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE			1 Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K., Arita,M., Musashino,H., Yuuki,H., Hara,H., Sugiyama,T., Irie,R., Otsuki.T., Sato.K., Wakamatsu.A., Ishii.S., Yamamoto,J., Iseno,Y., Kawai-Hio.Y., Saïto.K., Nishikawa.T., Kimura,K., Yamashita.H., Matsumoto,K., Nakamura.Y., Sekine.M., Kikuchi.H., Kanda.K., Wagatsuma,M., Murakawa,K., Kanehori.K., Sugiyama.A., Kawakami,B., Suzuki,X., Sugano,S., Nagahari,K., Masuko,Y., Nagai,X. and Isogai,I. NEDO human cDNA sequencing project
TITLE			NEDO human cDNA sequencing project
JOURNAL			Unpublished
REFERENCE			2 (bases 1 to 2343) Isogai,T. and Yamamoto,J. Direct Submission Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazuha-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:kgenomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); CDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES			Location/Qualifiers 1..2343 organism="Homo sapiens" mol_type="mRNA" db_xref="taxon:9606" cell_type="chondrocytes (HC)" notes="cloning vector: pME18FL3-primary culture, chondrocytes"
source			
ORIGIN			
Query Match			89.8%; Score 2240.8; DB 9; Length 2343;
Best local Similarity			99.9%; Pred. No. 0;
Matches 2253; Conservative			Mismatches 2; Indels 1; Gaps 1;
Y			224 TGCGCTGCGAGCCGTGGAGCGGCCTACTGCCCAATGGCCTCGCCCGCCGGT 283
D			89 TGCGCTGCGAGCCGTGGAGCGGCCTACTGCCCAATGGCCTCGCCCGCCGGT 148
			284 GC CGCG AG CT GG T GCG CG AG CC CG GT GCG GT GCT GCG ACT GAG C AT GAG CG AG G 343









RESULT 15	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	ORIGIN																																				
AK056274	2372 bp	linear	AK056274.1	GI:16551627	oligo capping; full insert sequence	Human	Homo sapiens	1	Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S., Ono, Y., Hotta, T., Hiraoka, S., Murakawa, K., Takiguchi, S., Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M., Yamashita, H., Chiba, Y., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.	NEDO human cDNA sequencing project	Unpublished	2	(bases 1 to 2372)	Isogai, T., Otsuki, T. and Sugiyama, T.	Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- and 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.	Location/Qualifiers 1..2372 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="NT2R12006445" /cell_line="NT2" /cell_type="teratocarcinoma" /clone_lib="NT2R12" /note="cloning vector: pMEJ8SFL3-mRNA from NT2 neuronal precursor cells treated 2-weeks mitotic inhibitor after 5-weeks retinoic acid (RA) induction.-majority NT2 neuron"	Query Match 88.6%; Score 2209.8; DB 9; Length 2372; Best Local Similarity 94.8%; Pred. No. 0; Matches 2349; Conservative 0; Mismatches 2; Indels 128; Gaps 1;	1	CGCTGAGTACAGCGCTTCTGCTGGATCCACAGCTCGCGCGCTGCTGCTGCGC	60		22	GGCGTGGAGATCAGCGCTTCTGCTGGATCCACAGCTTGGCGCGCTGCTGCTGCGC	81		61	CCATCCCTCGCGCCAGCGCTGCCAAGCAGCGTCCCGCGTTCGAGCGTTCATGCGCG	120		92	CCATCCTTGGCGCCAGCGTCCCAAGCAGCGTCCCGCGTTCGAGCGTTCATGCGCG	141		121	GGCGCAGCCAGCTTGGCGCGCTCGCTGACTCTGCTGGTGTGCTCGCGCGCGCGCG	180		142	GGCGCAGCCAGCTTGGCGCGCTCGCTGACTCTGCTGGTGTGCTCGCGCGCGCGCG	201		181	GTGGCGCGGCTGGCGCGAGCTCGCGGGGCTTGGGTCCTCGCTCGCGCGCGCTGC	240		202	GTGGCGCGGCTGGCGCGAGCTCGCGGGGCTTGGGTCCTCGCTCGCGCGCGCTGC	240		241	GACGCGCGTGCATGGCGCGCTTCTGCTGGATCCACAGCTTGGCGCGCTTGGCGCG	300	

QY 1321 CCACTGACTTTGTGACTTAGGGGGTGTGTGCTATGTAGAGAACAGCGTTTCAACCCCA 1380  
 Db 1214 CCACTGACTTTGTGACTTAGGGGGTGTGTGCTATGTAGAGAACAGCGTTTCAACCCCA 1273  
 QY 1381 CTCGCCGTACAGTGGCCACAGCGTTTATCGAGATAGGAAACCTTTAAACCCCGTCTAT 1440  
 Db 1274 CTCGCCGTACAGTGGCCACAGCGTTTATCGAGATAGGAAACCTTTAAACCCCGTCTAT 1333  
 QY 1441 CCGGACATCCCAACGCGATGCTCTGAGAGTCCACAGCTTCTGTGGTGTCTTCTGAAAC 1500  
 Db 1334 CCGGACATCCCAACGCGATGCTCTGAGAGTCCACAGCTTCTGTGGTGTCTTCTGAAAC 1393  
 QY 1501 AAGGGGTGATCCCTCAACCAAGAGATGTTTATGCTTCAAGTGACCTGTACTGCTT 1560  
 Db 1394 AAGGGGTGATCCCTCAACCAAGAGATGTTTATGCTTCAAGTGACCTGTACTGCTT 1453  
 QY 1561 GGGGACTATTGGAGAAATAAGGTGGAGTCTTACTTGTAAAAAATATGTATCTAAGAA 1620  
 Db 1454 GGGGACTATTGGAGAAATAAGGTGGAGTCTTACTTGTAAAAAATATGTATCTAAGAA 1513  
 QY 1621 TGTTCAGGCACTCTGGGAACCTATTAAGGCAAGGTATTTGGGGCCCTCTCTTCAGGAA 1680  
 Db 1514 TGTTCAGGCACTCTGGGAACCTATTAAGGCAAGGTATTTGGGGCCCTCTCTTCAGGAA 1573  
 QY 1681 TCTTCTGAAGACATGCCCGAGTCGAAGGCCAGGATGCTTTTGTGGGGCCCGTGGG 1740  
 Db 1574 TCTTCTGAAGACATGCCCGAGTCGAAGGCCAGGATGCTTTTGTGGGGCCCGTGGG 1633  
 QY 1741 GTAGGAGGACAGAGACAGAGGAGTCAAGCTCCACATTCAGAGGCAATCAAGTAAT 1800  
 Db 1634 GTAGGAGGACAGAGACAGAGGAGTCAAGCTCCACATTCAGAGGCAATCAAGTAAT 1693  
 QY 1801 GGCACAACTTCTCGGATGCTCGAGAAATAGTGTGTTGTAGTTCAAACACTCAAGACGA 1860  
 Db 1694 GGCACAACTTCTCGGATGCTCGAGAAATAGTGTGTTGTAGTTCAAACACTCAAGACGA 1753  
 QY 1861 AGCTTAATTTCTAGGATAAGCTCTTTAAAGGCAAGCTTATTTTCACTCTCATCTTTT 1920  
 Db 1754 AGCTTAATTTCTAGGATAAGCTCTTTAAAGGCAAGCTTATTTTCACTCTCATCTTTT 1813  
 QY 1921 GTCTCCTTAGCACATGTAAGAAAGATAGTAATATCAGACAGGAAGGAGGCT 1980  
 Db 1814 GTCTCCTTAGCACATGTAAGAAAGATAGTAATATCAGACAGGAAGGAGGCT 1873  
 QY 1981 TGCTGGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTTGTGAAAC 2040  
 Db 1874 TGCTGGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTTGTGAAAC 1933  
 QY 2041 TTAGAGTCAATCTCATGCTTTTCTTTATATTCACACATATATGCAGAGAGATATGTTTC 2100  
 Db 1934 TTAGAGTCAATCTCATGCTTTTCTTTATATTCACACATATATGCAGAGAGATATGTTTC 1993  
 QY 2101 TTGTTAAACATTGTATACACATAGCCCCAAATATAGTAAGATCTATCTAGATAATCCTA 2160  
 Db 1994 TTGTTAAACATTGTATACACATAGCCCCAAATATAGTAAGATCTATCTAGATAATCCTA 2053  
 QY 2161 GATGAATGTTAGAGATGCTATTTGTATACACTGTGGCCATGCTAGGAAAGGCTCA 2220  
 Db 2054 GATGAATGTTAGAGATGCTATATGATACAACTGTGGCCATGCTAGGAAAGGCTCA 2113  
 QY 2221 CGCCGAGAGACTGGGCTGCTCTCCGGAGGCCAAACCCAGAAAGGCTCGCCAAAGTCAGG 2280  
 Db 2114 CGCCGAGAGACTGGGCTGCTCTCCGGAGGCCAAACCCAGAAAGGCTCGCCAAAGTCAGG 2173  
 QY 2281 CTCAGGAGACTCTGCCCTGTGTGAGACCTCGGTGTGGACACAGCTGCAATAGAGCTTC 2340  
 Db 2174 CTCAGGAGACTCTGCCCTGTGTGAGACCTCGGTGTGGACACAGCTGCAATAGAGCTTC 2233  
 QY 2341 CTTGAAACAGAGGGGTCTCAAGACATTTCTGCCTACCTATAGCTTTTCTTTATTTTTT 2400  
 Db 2234 CTTGAAACAGAGGGGTCTCAAGACATTTCTGCCTACCTATAGCTTTTCTTTATTTTTT 2293  
 QY 2401 AACTTTTTGGGGGAAAGATATTTTGTGAGAGTTTGTCTTGGCAATGTATTTTATAATAGT 2460

Db 2294 AACTTTTTGGGGGAAAGATATTTTGTGCAATGTATTTATAAATAGT 2353  
 QY 2461 AAATAAAGTTTTTACCATT 2479  
 Db 2354 AAATAAAGTTTTTACCATT 2372

Search completed: March 11, 2004, 03:35:45  
 Job time : 6369.58 secs